

QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLECECEKPLEKSHCIAEVENDEMPA 300
 Db 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLECECEKPLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYVYARRHPDYSVLLRLAKTYETTTLEKC 360
 Db 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYVYARRHPDYSVLLRLAKTYETTTLEKC 360
 QY 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFQELGEYKFQNALVRYTKVPQVST 420
 Db 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFQELGEYKFQNALVRYTKVPQVST 420
 QY 421 PTLVEVSRNLGVKSGCKCKHPEAKRVPCEADYLSVVLNQLCVLHKTVPVSRVTKCCTES 480
 Db 421 PTLVEVSRNLGVKSGCKCKHPEAKRVPCEADYLSVVLNQLCVLHKTVPVSRVTKCCTES 480
 QY 481 LVNRRPCFSALEVDYTVPKFNAETFTFHADICTLSEKERQIKKOTALVELVHKPKAT 540
 Db 481 LVNRRPCFSALEVDYTVPKFNAETFTFHADICTLSEKERQIKKOTALVELVHKPKAT 540
 QY 541 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEKGKLVAAASQAALGL 585
 Db 541 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEKGKLVAAASQAALGL 585

RESULT 16

ABG33321
 ID ABG63321 standard; protein; 585 AA.
 XX AC ABG63321;
 XX DT 27-AUG-2002 (first entry)
 XX DE Human serum albumin (HSA) protein.
 XX KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
 KW human serum albumin; HSA; cancer; reproductive disorder;
 KW digestive disorder; immune disorder; endocrine disorder;
 KW haematopoietic disorder; neural disorder; connective disorder;
 KW cytostatic; antiinfertility; antiinflammatory; antiulcer;
 KW immunomodulator; anti-HIV; antidiabetic; haemostatic; neuroleptic;
 KW neuroprotective; antiparkinsonian; antimicrobial; osteopathic; antiarthritic.
 XX OS Homo sapiens.
 XX PN WO200177137-A1.
 XX PD 18-OCT-2001.
 XX PF 12-APR-2001; 2001WO-US011988.
 XX PR 12-APR-2000; 2000US-0229358P.
 XX PR 25-APR-2000; 2000US-0199384P.
 XX PR 21-DEC-2000; 2000US-0256931P.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Rosen CA, Haseltine WA;
 XX WPI: 2002-010886/01.
 XX DR N-PSDB; ABK93280.
 XX PT New fusion protein for treating disease e.g. diabetes comprises an
 XX albumin fused to a therapeutic protein.
 XX PS Claim 1; Fig 15; 2102pp; English.
 XX CC The present invention relates to albumin fusion proteins comprising a
 CC therapeutic protein X and human albumin (HA), also known as human serum
 CC albumin, HSA). The proteins are useful for treating a disease or disorder
 CC that may be modulated by therapeutic protein X. The albumin extends the
 CC shelf-life of protein X, and may increase its biological in vitro/in vivo

CC activity. The protein is useful for treating and diagnosing disorders
 CC such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's
 CC disease, ulcerative colitis), immune disorders (e.g. acquired
 CC immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),
 CC haematopoietic disorders, neural disorders (e.g. Alzheimer's,
 CC Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,
 CC schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).
 CC The present sequence represents HSA (HA) protein
 XX Sequence 585 AA;
 SQ
 Query Match 100.0%; Score 585; DB 5; Length 585;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHKEVAHRFMDLGEENFKALVLIAPAYLIQQCPFFEDHVKLVNEVTEFAKTCVADSEAE 60
 Db 1 DAHKEVAHRFMDLGEENFKALVLIAPAYLIQQCPFFEDHVKLVNEVTEFAKTCVADSEAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKGEPERNECFLOHKDDNPNLPLVPEV 120
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKGEPERNECFLOHKDDNPNLPLVPEV 120
 QY 121 DVMCTAFHDNEETFLKYLVIETARRHPYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
 Db 121 DVMCTAFHDNEETFLKYLVIETARRHPYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
 QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRFPAEFAEYVKLVDTLTK 240
 Db 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRFPAEFAEYVKLVDTLTK 240
 QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLECECEKPLEKSHCIAEVENDEMPA 300
 Db 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLECECEKPLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYVYARRHPDYSVLLRLAKTYETTTLEKC 360
 Db 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYVYARRHPDYSVLLRLAKTYETTTLEKC 360
 QY 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFQELGEYKFQNALVRYTKVPQVST 420
 Db 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFQELGEYKFQNALVRYTKVPQVST 420
 QY 421 PTLVEVSRNLGVKSGCKCKHPEAKRVPCEADYLSVVLNQLCVLHKTVPVSRVTKCCTES 480
 Db 421 PTLVEVSRNLGVKSGCKCKHPEAKRVPCEADYLSVVLNQLCVLHKTVPVSRVTKCCTES 480
 QY 481 LVNRRPCFSALEVDYTVPKFNAETFTFHADICTLSEKERQIKKOTALVELVHKPKAT 540
 Db 481 LVNRRPCFSALEVDYTVPKFNAETFTFHADICTLSEKERQIKKOTALVELVHKPKAT 540
 QY 541 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEKGKLVAAASQAALGL 585
 Db 541 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEKGKLVAAASQAALGL 585

RESULT 17

ABG33847
 ID ABG33847 standard; protein; 585 AA.
 XX AC ABG33847;
 XX DT 15-JUL-2002 (first entry)
 XX DE Human B Lymphocyte Stimulator (BlyS) protein #3.
 XX KW B Lymphocyte Stimulator protein; B Lymphocyte Stimulator binding peptide;
 KW BlyS; biological fluid; serum; plasma; lymph; blood; urine; spinal fluid;
 KW synovial fluid; saliva; mucus; human.
 XX OS Homo sapiens.
 XX PN WO200216412-A2.

XX PD 28-FEB-2002.

XX PF 17-AUG-2001; 2001WO-US025891.

XX PR 18-AUG-2000; 2000US-0226489P.

XX PA (DYAX-) DYAX CORP.

XX PI Beltzer JP, Potter MD, Fleming TJ, Ladner RC;

XX DR WPI; 2002-351647/38.

XX DD

XX PT New B-lymphocyte stimulator binding polypeptide useful in detecting or

XX PT isolating Blys or Blys-like polypeptide comprises a specified amino acid

XX PT sequence.

XX PS Disclosure; Page 261-264; 269pp; English.

XX CC The invention relates to a B Lymphocyte Stimulator (Blys) binding

XX CC polypeptide. Blys binding peptides bind Blys or Blys-like proteins

XX CC reversibly or irreversibly. The binding peptides are used in detection,

XX CC isolation and/or purification of Blys in a solution such as water or a

XX CC buffer solution, as well as any fluid and/or cell obtained from an

XX CC individual biological fluid, body tissue, body cell, cell line, tissue

XX CC culture or other source containing Blys or Blys-like polypeptides. The

XX CC biological fluids include sera, plasma, lymph, blood, blood fraction,

XX CC urine, synovial fluid, spinal fluid, saliva and mucus. Sequences

XX CC ABG33576, ABG33577 and ABG33847 represent human B Lymphocyte Stimulator

XX CC proteins

XX PA

XX DR

XX SQ Sequence 585 AA;

Query Match 100.0%; Score 585; DB 5; Length 585;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHFRKDLGEENFKALVLIAPVLOQCPEDHVKLVNEVTEFAKTCVADESAE 60

DB 1 DAHKEVAHFRKDLGEENFKALVLIAPVLOQCPEDHVKLVNEVTEFAKTCVADESAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

QY 121 DVMTAFHDNEETFLKKYLIETARRHPYFVAPPELLFFAKRYKAAFTCCQAADKAACLLP 180

DB 121 DVMTAFHDNEETFLKKYLIETARRHPYFVAPPELLFFAKRYKAAFTCCQAADKAACLLP 180

QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFAEVSCLVTDLT 240

DB 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFAEVSCLVTDLT 240

QY 241 VHTECCHGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300

DB 241 VHTECCHGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300

QY 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360

DB 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360

QY 361 CAADPHCEYAKVDFEFPKLVEEPQNLIKONCELFEGQYKFNQALLVRYTKVPQVST 420

DB 361 CAADPHCEYAKVDFEFPKLVEEPQNLIKONCELFEGQYKFNQALLVRYTKVPQVST 420

QY 421 PTLVEVSRLGKVGSKCKCHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRYTKCCTES 480

DB 421 PTLVEVSRLGKVGSKCKCHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRYTKCCTES 480

QY 481 LVNRRPCSALEVEDTVVPKEFNAETFFHADICTLSEKERQIKKQATLVELVGHKPKAT 540

DB 481 LVNRRPCSALEVEDTVVPKEFNAETFFHADICTLSEKERQIKKQATLVELVGHKPKAT 540

QY 541 KEQLKAVNDDFAAFVEKCKKADDKETCFABEGKLVAAASQAALGL 585

DB 541 KEQLKAVNDDFAAFVEKCKKADDKETCFABEGKLVAAASQAALGL 585

RESULT 18

ABG71291

ID ABG71291 standard; protein; 585 AA.

XX AC ABG71291;

XX DT 08-JAN-2003 (first entry)

XX DE Glycosylated protein determination associated protein.

XX DE Standard substance; accuracy control substance; glycosylated protein;

XX KW glycosylated albumin; fructoseamine; diabetes; antidiabetic.

XX OS Unidentified.

XX PN JP2002243731-A.

XX PD 28-AUG-2002.

XX PF 21-FEB-2001; 2001JP-00045085.

XX PR 21-FEB-2001; 2001JP-00045085.

XX XX (KOKU-) KOKUSAI SHIYAKU KK.

XX PA (YOSH) YOSHITOMI PHARM IND KK.

XX DR WPI; 2002-744850/81.

XX PT A standard substance for determination of glycosylated protein including

XX PT glycosylated albumin and fructoseamine, used in diagnosis of diabetes.

XX PS Disclosure; Page 4; 6pp; Japanese.

XX CC The present invention relates to a new standard and accuracy control

XX CC substance for determination of glycosylated protein. The invention is

XX CC useful for determination of glycosylated protein in the diagnosis of

XX CC diabetes. Glycosylated albumin and fructoseamine provide favourable

XX CC dilution linearity. The present amino acid sequence represents the

XX CC glycosylated protein determination associated protein as described in the

XX CC invention

XX SQ Sequence 585 AA;

Query Match 100.0%; Score 585; DB 5; Length 585;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHFRKDLGEENFKALVLIAPVLOQCPEDHVKLVNEVTEFAKTCVADESAE 60

DB 1 DAHKEVAHFRKDLGEENFKALVLIAPVLOQCPEDHVKLVNEVTEFAKTCVADESAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

QY 121 DVMTAFHDNEETFLKKYLIETARRHPYFVAPPELLFFAKRYKAAFTCCQAADKAACLLP 180

DB 121 DVMTAFHDNEETFLKKYLIETARRHPYFVAPPELLFFAKRYKAAFTCCQAADKAACLLP 180

QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFAEVSCLVTDLT 240

DB 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFAEVSCLVTDLT 240

QY 241 VHTECCHGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300

DB 241 VHTECCHGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300

QY 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360

Db 301 DFLPSLAADPVESKDYCKNYAEAKVFLGMFLYVARRHPDYSVLLRLAKTYETTLK 360
QY 361 CAAADPHECYAKVDFEFPKPLVBEPPONLIKONCELPQOLGEYKFNALLVRYTKVPQVST 420
Db 361 CAAADPHECYAKVDFEFPKPLVBEPPONLIKONCELPQOLGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEYDLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEYDLSVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQALVELVGHKPKAT 540
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQALVELVGHKPKAT 540
QY 541 KEQLKAVMDDFAAVFEKCKCKADDDKTCFAEGKKLVAASQAALGL 585
Db 541 KEQLKAVMDDFAAVFEKCKCKADDDKTCFAEGKKLVAASQAALGL 585

RESULT 19
ABR55695
ID ABR55695 standard; protein; 585 AA.
XX ABR55695;
XX
DT 18-AUG-2003 (first entry)
XX Human albumin (HA) mature polypeptide.
DE
XX
KW Albumin; HA; cytostatic; antibacterial; virucide; fungicide; anti-HIV;
KW antiasthmatic; osteopathic; antiarthritic; antiinflammatory; nootropic;
KW neuroprotective; anti-thyroid; anti-ulcer; hepatotropic; vulnerary;
KW protein therapy.
XX
OS Homo sapiens.
XX
XX WO2003030821-A2.
XX
XX 17-APR-2003.
XX
XX 04-OCT-2002; 2002WO-US031794.
XX
XX 05-OCT-2001; 2001US-0327281P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Haseltine WA;
XX
XX WPI; 2003-468174/44.
XX
XX N-PSDB; ACC78574.
XX
XX New albumin fusion protein comprising a therapeutic protein.X, and
XX albumin, its variant or fragment, useful for treating a cancer, AIDS
XX asthma, leukemia, sepsis, endometriosis, osteoporosis, atherosclerosis,
XX autism, or emphysema.
XX
XX Claim 1; Fig 15A-D; 455pp; English.

PS
XX
XX The invention relates to an albumin fusion protein comprising a
XX therapeutic protein.X, and albumin, its variant or fragment. The albumin
XX fusion protein has the formula R1-L-R2; R2-L-R1; or R1-L-R2-L-R1 where R1
XX is therapeutic protein.X or fragment, L is a peptide linker and R2 is
XX albumin. The albumin fusion protein is useful for treating a disease or
XX disorder that is modulated by therapeutic protein.X (claimed), such as
XX cancer; infections (bacterial, viral, fungal, parasitic); or immune
XX (AIDS, asthma); hematopoietic (leukemia, sepsis); reproductive (cystic
XX fibrosis, endometriosis); musculoskeletal (osteoporosis, osteoarthritis);
XX cardiovascular (congestive heart failure, atherosclerosis); neural/
XX sensory (ataxia, attention deficit disorders, autism); respiratory
XX (emphysema, bronchitis); endocrine (goiter, glomerulonephritis);
XX digestive (ulcer, cirrhosis); or connective/epithelial (lupus, keloids)
XX disorders. The present sequence represents a human albumin (HA) mature

CC polypeptide
XX Sequence 585 AA;
SQ
Query Match 100.0%; Score 585; DB 6; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHREKDI GENFKALVLI AFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
Db 1 DAHSEVAHREKDI GENFKALVLI AFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKYLBIARRHPFYFAPPELLFFAKRYKAAPTECCCAADKAACLLP 180
Db 121 DVMCTAFHDNEETFLKYLBIARRHPFYFAPPELLFFAKRYKAAPTECCCAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORPPKAEFAFVSKLVTDLTK 240
Db 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORPPKAEFAFVSKLVTDLTK 240
QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300
Db 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300
QY 301 DPLSLAADPVESKDYCKNYAEAKVFLGMFLYVARRHPDYSVLLRLAKTYETTLK 360
Db 301 DPLSLAADPVESKDYCKNYAEAKVFLGMFLYVARRHPDYSVLLRLAKTYETTLK 360
QY 361 CAAADPHECYAKVDFEFPKPLVBEPPONLIKONCELPQOLGEYKFNALLVRYTKVPQVST 420
Db 361 CAAADPHECYAKVDFEFPKPLVBEPPONLIKONCELPQOLGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEYDLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEYDLSVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQALVELVGHKPKAT 540
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQALVELVGHKPKAT 540
QY 541 KEQLKAVMDDFAAVFEKCKCKADDDKTCFAEGKKLVAASQAALGL 585
Db 541 KEQLKAVMDDFAAVFEKCKCKADDDKTCFAEGKKLVAASQAALGL 585

RESULT 20
ABR42606
ID ABR42606 standard; protein; 585 AA.

XX ABR42606;
XX
DT 26-AUG-2003 (first entry)
XX
DE Human serum albumin.
XX
XX Human; serum albumin; abrogen; angiogenesis; inhibitor;
XX urokinase plasminogen activator; tumour; metastasis; cytostatic;
XX gene therapy.
XX
XX Homo sapiens.
XX
XX WO2003042354-A2.
XX
XX 22-MAY-2003.
XX
XX 04-SEP-2002; 2002WO-US027885.
XX
XX 04-SEP-2001; 2001US-0316300P.
XX

PA (AVET) AVENTIS PHARM INC.
XX Nesbit M, Fong TC, Brockstedt D;
PI
XX WPI; 2003-449566/42.
DR
XX New abrogen polypeptide, useful for treating an angiogenesis related
PT diseases e.g. tumor metastasis.
XX
PS Disclosure; Page 26; 95pp; English.
XX
CC The present sequence is the protein sequence of human serum albumin
CC (HSA). The invention provides novel abrogen polypeptides (see ABR42599-
CC 602) that include a kringle domain, e.g. from urokinase plasminogen
CC activator kringle domain. Abrogens are potent inhibitors of endothelial
CC proliferation and angiogenesis, and have been shown to be capable of
CC inhibiting or reducing cell proliferation induced by both basic
CC fibroblast growth factor and vascular endothelial growth factor in a
CC specific endothelial cell proliferation assay. The abrogen may be coupled
CC to an N-terminal interleukin-2 signal peptide and to a C-terminal
CC stabilising molecule such as HSA (see ABR42608-13). Abrogen polypeptides
CC and polynucleotides are used in claimed methods of treating an
CC angiogenesis-related disease or disorder, e.g. tumour metastasis
XX
XX Sequence 585 AA;
SQ
Query Match 100.0%; Score 585; DB 7; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DAHKSEVAHRFKDLGSENFKALVLIAPAFYLOQCPEDHVKLVNEVTFEAKTCVADESAAE 60
Db 1 DAHKSEVAHRFKDLGSENFKALVLIAPAFYLOQCPEDHVKLVNEVTFEAKTCVADESAAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLOHKDNPMLPRLVREPV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLOHKDNPMLPRLVREPV 120
Qy 121 DVMCTAFHDNEETFLKKYLYEIAARRPYFYAPPELLFFAKRYKAAATECCQAADKAACLLP 180
Db 121 DVMCTAFHDNEETFLKKYLYEIAARRPYFYAPPELLFFAKRYKAAATECCQAADKAACLLP 180
Qy 181 KLDELDEGKASSAKQRLKCAISLQKGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
Db 181 KLDELDEGKASSAKQRLKCAISLQKGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
Qy 241 VHTCCCHGDLLECADRRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDENPA 300
Db 241 VHTCCCHGDLLECADRRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDENPA 300
Qy 301 DLPSLAADFVSKDVCNVAEAKDVLGMFLYEAARRHPDYSVLLRLAKTYETTLK 360
Db 301 DLPSLAADFVSKDVCNVAEAKDVLGMFLYEAARRHPDYSVLLRLAKTYETTLK 360
Qy 361 CAADPHCEYAKVDFEFPKLVBPQNLIKONCELFEPQLGEYKFNALLVRYTKVPQYST 420
Db 361 CAADPHCEYAKVDFEFPKLVBPQNLIKONCELFEPQLGEYKFNALLVRYTKVPQYST 420
Qy 421 PTLVEVSRLNGKSGCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
Db 421 PTLVEVSRLNGKSGCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
Qy 481 LVNRRPCFSALVDETVVPKFNARFTTHADICTLSEKEROIKKOTALVELVKHKPKAT 540
Db 481 LVNRRPCFSALVDETVVPKFNARFTTHADICTLSEKEROIKKOTALVELVKHKPKAT 540
Qy 541 KEQLKAVMDDFAAFEVCKCKADDKETCTCFABEGKKLVAASQAALGL 585
Db 541 KEQLKAVMDDFAAFEVCKCKADDKETCTCFABEGKKLVAASQAALGL 585
RESULT 21
ADC16767

ID ADC16767 standard; protein; 585 AA.
XX
AC ADC16767;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human serum albumin protein used to construct fusion proteins.
XX
KW human; anti-retroviral; T-20; T-1249; 5-Helix; cyanovirin-N; env; gp41;
KW anti-HIV; vaccine; albumin fusion protein; HIV fusion inhibiting peptide;
KW serum albumin.
XX
OS Homo sapiens.
XX
XX WO2003066078-A1.
XX 14-AUG-2003.
XX
XX 07-FEB-2003; 2003WO-IB0000434.
XX
XX 07-FEB-2002; 2002US-0355547P.
XX
XX (AVET) AVENTIS BEHRING GMBH.
XX (DELZ) DELTA BIOTECHNOLOGY LTD.
XX
XX Hauser H, Welmer T, Sleep D;
XX
XX WPI; 2003-731478/59.
DR N-PSDB; ADC16766.
XX
XX New albumin fusion protein comprising a human immunodeficiency virus
PT (HIV) fusion inhibiting peptide and an albumin having an albumin
PT activity, useful for treating a disease or disorder, e.g. HIV infection.
XX
PS Disclosure; Fig 14; 105pp; English.
XX
CC This invention relates to novel albumin fusion proteins comprising a
CC human immunodeficiency virus (HIV) fusion inhibiting peptide, which
CC exhibit anti-retroviral activity. Specifically, it refers to inhibitory
CC peptides including T-20, T-1249, 5-Helix or cyanovirin-N that bind the
CC HIV env protein, or derivatives thereof such as the HIV gp41 protein.
CC Furthermore, the albumin activity has the ability to prolong the in vivo
CC half-life of these HIV fusion inhibiting peptides. Accordingly, the
CC present invention describes fusion proteins that neutralise HIV in a host
CC by raising an immune response and also antibodies that inhibit viral
CC infection of uninfected cells. In this way, a method exists to prevent,
CC treat or ameliorate HIV infection and/or a disease caused by HIV
CC infection. As such, these composition have been described as having anti-
CC HIV activity and can be used towards the production of a vaccine. This
CC polypeptide sequence is the human serum albumin protein used to construct
CC fusion proteins of the invention.
XX
XX Sequence 585 AA;
SQ
Query Match 100.0%; Score 585; DB 7; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DAHKSEVAHRFKDLGSENFKALVLIAPAFYLOQCPEDHVKLVNEVTFEAKTCVADESAAE 60
Db 1 DAHKSEVAHRFKDLGSENFKALVLIAPAFYLOQCPEDHVKLVNEVTFEAKTCVADESAAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLOHKDNPMLPRLVREPV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLOHKDNPMLPRLVREPV 120
Qy 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAATECCQAADKAACLLP 180
Db 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAATECCQAADKAACLLP 180
Qy 181 KLDELDEGKASSAKQRLKCAISLQKGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
Db 181 KLDELDEGKASSAKQRLKCAISLQKGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK 240

QY 241 VHTCCGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPKLLEKSHCHIAEVENDEMPA 300
 DB |||||
 QY 241 VHTCCGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPKLLEKSHCHIAEVENDEMPA 300
 DB |||||
 QY 301 DLPSLAADFVSKOVCKNYAEAKOVFLGMFLYVYARHPDYSVLLRLAKTVETTLK 360
 DB |||||
 QY 301 DLPSLAADFVSKOVCKNYAEAKOVFLGMFLYVYARHPDYSVLLRLAKTVETTLK 360
 DB |||||
 QY 361 CAADPHCYAKVDFEFKPLVEEPQNLKQNCLEFEQLGKEYKFNALLVRYTKVPQVST 420
 DB |||||
 QY 361 CAADPHCYAKVDFEFKPLVEEPQNLKQNCLEFEQLGKEYKFNALLVRYTKVPQVST 420
 DB |||||
 QY 421 PTLVSVNRLGKVGSKCKHPEAKRMPCAEDYLSVNLQNLVILHEKTPVSDRVTKCTES 480
 DB |||||
 QY 421 PTLVSVNRLGKVGSKCKHPEAKRMPCAEDYLSVNLQNLVILHEKTPVSDRVTKCTES 480
 DB |||||
 QY 481 LVNRRPCFSALEVDYTVYKPFNAETFTFHADICTLSEKERQIKQATLVELVKHPKAT 540
 DB |||||
 QY 481 LVNRRPCFSALEVDYTVYKPFNAETFTFHADICTLSEKERQIKQATLVELVKHPKAT 540
 DB |||||
 QY 541 KEQLKAVMDDFAAVYKCKKADDETCFAEKGKLVAAQALGL 585
 DB |||||
 QY 541 KEQLKAVMDDFAAVYKCKKADDETCFAEKGKLVAAQALGL 585
 DB |||||

RESULT 22
 ADD06469
 ID ADD06469 standard; protein; 585 AA.
 AC ADD06469;
 XX ADD06469;
 DT 01-JAN-2004 (first entry)
 XX Human serum albumin protein SEQ ID NO:5.
 DE human; chemokine betal; Ckb1; anti-HIV; neuroprotective; antithyroid;
 KW antithyroid; antirheumatic; immunosuppressive; nontropic;
 KW antirheumatic; antiasthmatic; antiallergic; osteopathic;
 KW nephrotrophic; tuberculostatic; virucide; antiatherosclerotic;
 KW antitubercular; infection; HIV; immune disorder; haematopoietic disorder;
 KW autoimmune disorder; multiple sclerosis; Grave's disease; arthritis;
 KW rheumatoid arthritis; transplant rejection; neurodegenerative disorder;
 KW Alzheimer's disease; inflammatory disease; asthma; allergic disorder;
 KW inflammatory bowel disease; osteoarthritis; colitis;
 KW inflammatory kidney disease; glomerulonephritis; infectious disease;
 KW tuberculosis; hepatitis infection; herpes viral infection;
 KW viral infection; proliferative disorder; atherosclerosis;
 KW human serum albumin; HSA.
 XX Homo sapiens.
 OS
 XX
 XX WO200297038-A2.
 PN
 XX
 XX
 XX 05-DEC-2002.
 PD
 XX
 XX 24-MAY-2002; 2002WO-US016525.
 PF
 XX
 XX 25-MAY-2001; 2001US-0293212P.
 PR
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX
 XX Bell A, Ruben SM;
 PI
 XX
 XX WPI; 2003-140456/13.
 DR
 XX
 XX N-PSDB; ADD06468.
 DR
 XX
 XX Novel human chemokine betal protein comprising deletion in amino acids
 PT from amino and/or carboxy terminus, and is a fusion protein further
 PT comprising human serum albumin, is useful for treating multiple
 PT sclerosis, asthma.
 XX
 XX Disclosure; SEQ ID NO 5; 423pp; English.
 PS

XX The present invention describes a human chemokine betal (Ckb1) protein
 CC (I) comprising a deletion in amino acid residues from the amino terminus
 CC and/or carboxy terminus of the 93 residue amino acid sequence (S1, see
 CC ADD06466). (I) has anti-HIV, neuroprotective, antithyroid, antiasthmatic,
 CC antirheumatic, immunosuppressive, nontropic, antiallergic, antiatherosclerotic,
 CC antitubercular, antitubercular, osteopathic, nephrotrophic, tuberculostatic,
 CC virucide, antiatherosclerotic and antimicrobial activities. (I) is useful
 CC for preventing infection, preferably viral (human immunodeficiency virus
 CC (HIV)) infection, in a cell, by contacting the cell with (I). (I) is also
 CC useful for treating a disease, such as HIV infection or immune disorders,
 CC haematopoietic disorders, autoimmune disorders, multiple sclerosis,
 CC Grave's disease, arthritis, rheumatoid arthritis, transplant rejection,
 CC neurodegenerative disorders, Alzheimer's disease, inflammatory disease,
 CC asthma, allergic disorders, inflammatory bowel disease, osteoarthritis,
 CC disease, inflammatory kidney diseases, glomerulonephritis, infectious,
 CC disease, tuberculosis, hepatitis infections, herpes viral infection,
 CC individual. (I) inhibits or abolishes the ability of HIV to bind to,
 CC enter into/fuse with (infect), and/or replicate in CCR5 expressing cells.
 CC (I) also acts as CCR5 agonists or antagonists, stimulate chemotaxis of
 CC CCR5-expressing cells, inhibit CCR5 ligand binding to a CCR5 molecule, or
 CC upregulate or downregulate CCR5 expression. (I) is useful as an
 CC immunological probe for the differential identification of the tissues or
 CC cell-types. (I)-human serum albumin (HSA) fusion proteins are useful for
 CC diagnosing, treating and preventing various disorders in mammals,
 CC preferably in humans. (I)-HSA fusion proteins are also useful as
 CC molecular weight markers on sodium dodecyl sulfate polyacrylamide gel
 CC electrophoresis techniques, for raising antibodies, and to test the
 CC biological activities of the Ckb1 protein. (I)-HSA fusion proteins are
 CC useful for screening for molecules that bind to the Ckb1 protein portion
 CC of the fusion protein. The present sequence is used in the
 CC exemplification of the present invention.
 XX Sequence 585 AA;
 SQ

Query Match 100.0%; Score 585; DB 7; Length 585;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHSEVAHRRFKDGEENFKALVIAFAQLQCCPFEDHVKLVNEVTEFAKTCVADES 60
 DB 1 DAHSEVAHRRFKDGEENFKALVIAFAQLQCCPFEDHVKLVNEVTEFAKTCVADES 60
 QY 61 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOKDDNPNLRLVRPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOKDDNPNLRLVRPEV 120
 QY 121 DVMTAFHDNEETFLKYLVEIARRHPVYFAPPELLFFAKRYKAFTCCQAADKAACLLP 180
 DB 121 DVMTAFHDNEETFLKYLVEIARRHPVYFAPPELLFFAKRYKAFTCCQAADKAACLLP 180
 QY 181 KLDELDEGKASSAKQRLKCSAQKQGERAFKAWAVARLSQRPFKAEFAEVSCLVTDLT 240
 DB 181 KLDELDEGKASSAKQRLKCSAQKQGERAFKAWAVARLSQRPFKAEFAEVSCLVTDLT 240
 QY 241 VHTCCGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPKLLEKSHCHIAEVENDEMPA 300
 DB 241 VHTCCGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPKLLEKSHCHIAEVENDEMPA 300
 QY 301 DLPSLAADFVSKOVCKNYAEAKOVFLGMFLYVYARHPDYSVLLRLAKTVETTLK 360
 DB 301 DLPSLAADFVSKOVCKNYAEAKOVFLGMFLYVYARHPDYSVLLRLAKTVETTLK 360
 QY 361 CAADPHCYAKVDFEFKPLVEEPQNLKQNCLEFEQLGKEYKFNALLVRYTKVPQVST 420
 DB 361 CAADPHCYAKVDFEFKPLVEEPQNLKQNCLEFEQLGKEYKFNALLVRYTKVPQVST 420
 QY 421 PTLVSVNRLGKVGSKCKHPEAKRMPCAEDYLSVNLQNLVILHEKTPVSDRVTKCTES 480
 DB 421 PTLVSVNRLGKVGSKCKHPEAKRMPCAEDYLSVNLQNLVILHEKTPVSDRVTKCTES 480
 QY 481 LVNRRPCFSALEVDYTVYKPFNAETFTFHADICTLSEKERQIKQATLVELVKHPKAT 540

DB 481 LVNRRPCFSALEVDYVPKFNATFTFHADICTLSEKQIKKQALVELVKKPKAT 540
541 KEQLKAVNDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
541 KEQLKAVNDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585

RESULT 23
ADD68016
ID ADD68016 standard; protein; 585 AA.
XX
AC ADD68016;
XX
DT 29-JAN-2004 (first entry)
XX
DE Mature form of human albumin (HA).
XX
KW Albumin fusion protein; therapeutic protein; HIV; osteoporosis; cancer;
KW wound; autoimmune disease; cardiovascular disease; hepatitis;
KW multiple sclerosis; psoriasis; graft-versus-host disease; stroke;
KW atherosclerosis; inflammation; anti-HIV; osteopathic; cytostatic;
KW vulnerrary; cardiant; hepatotropic; neuroprotective; antipsoriatic;
KW immunosuppressive; cerebroprotective; antiarteriosclerotic;
KW antiinflammatory; human; serum albumin; HA; HSA.
XX
OS Homo sapiens.
XX
DN US2003125247-A1.
XX
PD 03-JUL-2003.
XX
PF 12-APR-2001; 2001US-00833041.
XX
PR 12-APR-2000; 2000US-0229358P.
PR 25-APR-2000; 2000US-0199384P.
PR 21-DEC-2000; 2000US-0256931P.
XX
XX (ROSE/) ROSEN C A.
PA (HASE/) HASELTINE W A.
XX
XX Rosen CA, Haseltine WA;
XX
XX WPI; 2003-810996/76.
DR N-PSDB; ADD68015.
XX
XX New albumin fusion protein for diagnosing, preventing or treating
PT diseases (e.g. HIV, cancer, atherosclerosis or stroke) comprises a
PT therapeutic protein (e.g. cathepsin K or vascular endothelial growth
PT factor) and an albumin.
XX
PS Claim 1; SEQ ID NO 18; 180pp; English.
XX
XX The present invention relates to albumin fusion proteins comprising any
CC of the therapeutic proteins listed in the specification, or their
CC fragments or variants, and an albumin protein or its fragments or
CC variants. The invention also discloses pharmaceutical compositions
CC comprising the albumin fusion proteins, a kit comprising the albumin
CC fusion proteins, and methods for treating a disease or disorder in a
CC patient, that is modulated by the therapeutic protein or its fragment or
CC variant. The compositions and methods of the invention are useful in
CC diagnosing, preventing, treating or ameliorating diseases or disorders,
CC such as HIV, osteoporosis, cancer, wounds, autoimmune diseases,
CC cardiovascular diseases, hepatitis, multiple sclerosis, psoriasis, graft-
CC versus-host disease, stroke, atherosclerosis and inflammation. The
CC present sequence represents the mature form of human albumin
CC (HA, also known as human serum albumin, HSA).
XX
SQ Sequence 585 AA;

Query Match 100.0%; Score 585; DB 7; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRFKDGBENFKALVLIAPQYLOQCPEDHVKLVNEVTEFAKTCVADES 60
DB 1 DAHSEVAHRRFKDGBENFKALVLIAPQYLOQCPEDHVKLVNEVTEFAKTCVADES 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEBPERNECFLOHKKDDPNLPRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEBPERNECFLOHKKDDPNLPRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLIEIARRHPFYFAPPELLFFAKYKAAAFTECCQAADKAACLLP 180
DB 121 DVMCTAFHDNEETFLKKYLIEIARRHPFYFAPPELLFFAKYKAAAFTECCQAADKAACLLP 180
QY 181 KLDELRODEGKASSAKQRLKCAQLKQFGERAFKAWAVARLSQRFPKAEFAEVS KLVDLT 240
DB 181 KLDELRODEGKASSAKQRLKCAQLKQFGERAFKAWAVARLSQRFPKAEFAEVS KLVDLT 240
QY 241 VHTTECHGDLLECADDDRADLAKIYICENQDSISSKLKECKEPILLESKSHCIAEVENDEMPA 300
DB 241 VHTTECHGDLLECADDDRADLAKIYICENQDSISSKLKECKEPILLESKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLRLAKTYETTLEK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLRLAKTYETTLEK 360
QY 361 CAADPHECYAKVDFEKPVLVEEPQNLIKONCELFEOLGEYKFNALLVRYTKVQVST 420
DB 361 CAADPHECYAKVDFEKPVLVEEPQNLIKONCELFEOLGEYKFNALLVRYTKVQVST 420
QY 421 PTLVEVSRNLKGVSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVT KCCTES 480
DB 421 PTLVEVSRNLKGVSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVT KCCTES 480
QY 481 LVNRRPCFSALEVDYVPKFNATFTFHADICTLSEKQIKKQALVELVKKPKAT 540
DB 481 LVNRRPCFSALEVDYVPKFNATFTFHADICTLSEKQIKKQALVELVKKPKAT 540
QY 541 KEQLKAVNDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
DB 541 KEQLKAVNDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585

RESULT 24
AAB36542
ID AAB36542 standard; protein; 609 AA.
XX
AC AAB36542;
XX
DT 07-MAR-2001 (first entry)
XX
DE Recombinant human serum albumin (HSA) protein sequence #1.
XX
XX Human serum albumin; HSA.
XX
OS Homo sapiens.
XX
PN CN1266099-A.
XX
PD 13-SEP-2000.
XX
PF 04-MAR-1999; 99CN-00102745.
XX
PR 04-MAR-1999; 99CN-00102745.
XX
XX (MAQJ-) MAQJ BIOLOGICAL ENG SCI & TECH CO LTD.
XX
XX Liu Z;
XX
XX WPI; 2000-673206/56.
DR N-PSDB; AAC99308.
XX
XX Novel methods for chemical synthesis, expression and recombinant protein
PT production for human serum albumin reformed gene.

XX Example 1; Fig 3; 85pp; Chinese.
PS The present invention relates to two kinds of DNA sequences of coded
XX human serum albumin (HSA), i.e. design of structure-modified gene segment
CC of HSA and artificial total synthesis and a production process for large-
CC scale production of genetic recombinant HSA by using methanol, yeast and
CC engineering bacterium, and discovers that the structure-modified gene can
CC greatly increase the expression quantity of HSA. The production process
CC can make the structural gene of HSA obtain high-level expression under
CC the drive of promoter induced by methanol, and make the HSA expression
CC product secrete into the fermenting liquor culture medium, and provide
CC reliable test data for more large-scale pilot-amplification of gene
CC engineering HSA. The present sequence represents a recombinant HSA
CC protein from the present invention
XX
XX Sequence 609 AA;
Query Match 100.0%; Score 585; DB 3; Length 609;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRFKDLGEENFKALVLIATFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 25 DAHKSEVAHRFKDLGEENFKALVLIATFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHQDDNPNLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHQDDNPNLRLVRPEV 144
QY 121 DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPELFFAKRYKAAFTCCQAADKAAACLLP 180
Db 145 DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPELFFAKRYKAAFTCCQAADKAAACLLP 204
QY 181 KLDELDEGKASSAKQRLKCSLOKFGERAFKAWAVARLSQRPFKAFBVSQKLVTDLT 240
Db 205 KLDELDEGKASSAKQRLKCSLOKFGERAFKAWAVARLSQRPFKAFBVSQKLVTDLT 264
QY 241 VHTECHGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 265 VHTECHGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKQVCKNYAEAKOVFLGMFLYEYARRHPDYSVVLLRLAKTYETTLLEK 360
Db 325 DLPSLAADFVESKQVCKNYAEAKOVFLGMFLYEYARRHPDYSVVLLRLAKTYETTLLEK 384
QY 361 CAAADPHECYAKVDFEFPKPLVEEPQNLKQNCLEFEQLGEYKFNALLVRYTKVPQVST 420
Db 385 CAAADPHECYAKVDFEFPKPLVEEPQNLKQNCLEFEQLGEYKFNALLVRYTKVPQVST 444
QY 421 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVNLQCLVLEHKTVPVSDRVTKCCTES 480
Db 445 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVNLQCLVLEHKTVPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEQIKQTALVELVKHKPKAT 540
Db 505 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEQIKQTALVELVKHKPKAT 564
RESULT 25
ID AAY78147
XX AAY78147 standard; protein; 609 AA.
XX AC AAY78147;
XX DE 02-MAY-2000 (first entry)
XX Pre human serum albumin protein sequence.

Human serum albumin; HSA; pre-HSA; Pichia pastoris.
XX Homo sapiens.
XX CN1235981-A.
XX 24-NOV-1999.
XX 15-MAY-1998; 98CN-00110844.
XX 15-MAY-1998; 98CN-00110844.
XX (SHAN-) SHANGHAI INST BIOCHEMISTRY CHINESE ACAD.
XX Yuan Z, Qiu R, Wu X;
XX WPI; 2000-148246/14.
XX N-PSDB; AA288278.
XX Express and purification of human serum albumin in Pichia pasters -
XX comprises the construction of recombinated expression plasmid PPQO-HSA
XX (human serum albumin).
XX Disclosure; Fig 1; 18pp; Chinese.
XX The present invention describes an expression and purification method for
XX human serum albumin (HSA) in Pichia pastoris. The method involves the
XX construction of a recombination expression plasmid, PPQO-HSA, and the
XX high-efficiency separation and purification of expressed HSA. The method
XX can obtain sample purity higher than 99%. The present sequence represents
XX the pre HSA protein sequence
XX Sequence 609 AA;
Query Match 100.0%; Score 585; DB 3; Length 609;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRFKDLGEENFKALVLIATFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 25 DAHKSEVAHRFKDLGEENFKALVLIATFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHQDDNPNLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHQDDNPNLRLVRPEV 144
QY 121 DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPELFFAKRYKAAFTCCQAADKAAACLLP 180
Db 145 DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPELFFAKRYKAAFTCCQAADKAAACLLP 204
QY 181 KLDELDEGKASSAKQRLKCSLOKFGERAFKAWAVARLSQRPFKAFBVSQKLVTDLT 240
Db 205 KLDELDEGKASSAKQRLKCSLOKFGERAFKAWAVARLSQRPFKAFBVSQKLVTDLT 264
QY 241 VHTECHGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 265 VHTECHGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKQVCKNYAEAKOVFLGMFLYEYARRHPDYSVVLLRLAKTYETTLLEK 360
Db 325 DLPSLAADFVESKQVCKNYAEAKOVFLGMFLYEYARRHPDYSVVLLRLAKTYETTLLEK 384
QY 361 CAAADPHECYAKVDFEFPKPLVEEPQNLKQNCLEFEQLGEYKFNALLVRYTKVPQVST 420
Db 385 CAAADPHECYAKVDFEFPKPLVEEPQNLKQNCLEFEQLGEYKFNALLVRYTKVPQVST 444
QY 421 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVNLQCLVLEHKTVPVSDRVTKCCTES 480
Db 445 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVNLQCLVLEHKTVPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEQIKQTALVELVKHKPKAT 540
Db 505 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEQIKQTALVELVKHKPKAT 564

QY 541 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEKGKLVAAASQAALGL 585
DB 565 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEKGKLVAAASQAALGL 609

RESULT 26

AAB36549
ID AAB36549 standard; protein; 609 AA.

AC AAB36549;

XX 07-MAR-2001 (first entry)

DT 07-MAR-2001 (first entry)

XX Recombinant human serum albumin (HSA) protein sequence #1.

DE Recombinant human serum albumin; HSA.

XX Human serum albumin; HSA.

XX Homo sapiens.

XX CN1266100-A.

XX 13-SEP-2000.

XX 04-MAR-1999; 99CN-00102794.

XX 04-MAR-1999; 99CN-00102794.

XX (MAOJ-) MAOJI BIOLOGICAL ENG SCI & TECH CO LTD.

XX Liu Z;

XX WPI; 2000-673207/66.

XX Novel methods for the chemical synthesis, expression and recombinant protein production for human serum albumin reformed gene.

XX Example 1; Fig 3; 85pp; Chinese.

XX The present invention relates to two kinds of DNA sequences of coded human serum albumin (HSA), i.e. design of structure-modified gene segment of HSA and artificial total synthesis and a production process for large-scale production of genetic recombinant HSA by using methanol, yeast and engineering bacterium, and discovers that the structure-modified gene can greatly increase the expression quantity of HSA. The production process can make the structural gene of HSA obtain high-level expression under the drive of promoter induced by methanol, and make the HSA expression product secrete into the fermenting liquor culture medium, and provide reliable test data for more large-scale pilot-amplification of gene engineering HSA. The present sequence represents a recombinant HSA protein from the present invention

XX Sequence 609 AA;

QY Query Match 100.0%; Score 585; DB 3; Length 609;
DB Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDGLGKLVIAFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHKSEVAHRFKDGLGKLVIAFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKQBPNECEFLQHKDNPMLPRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKQBPNECEFLQHKDNPMLPRLVRPEV 144
QY 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 145 DVNCTAFHDNEETFLKKLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 204
QY 181 KLDELDEGKASAKQRLKCAKSLQKGERAFKAWAVARLSQPPKAEFAEVSKLVTDLTK 240
DB 205 KLDELDEGKASAKQRLKCAKSLQKGERAFKAWAVARLSQPPKAEFAEVSKLVTDLTK 264

QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
DB 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFYIYARRHPDYSVLLRLAKTYETTTLEKC 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVFGLMFYIYARRHPDYSVLLRLAKTYETTTLEKC 384
QY 361 CAAADPHECYAKVDFEFPKPLVEBPQNLIKQNCLEFQOLGEYKFNALLVRYTKVPQVST 420
DB 385 CAAADPHECYAKVDFEFPKPLVEBPQNLIKQNCLEFQOLGEYKFNALLVRYTKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLSVNLNQLCVLHEKTPVSDRVTCKCTES 480
DB 445 PTLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLSVNLNQLCVLHEKTPVSDRVTCKCTES 504
QY 481 LVNRRPCPSALEVDETVVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVKHKPKAT 540
DB 505 LVNRRPCPSALEVDETVVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVKHKPKAT 564
QY 541 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEKGKLVAAASQAALGL 585
DB 565 KEQLKAVMDDDFAAFVEKCKKADDDKTCFAEKGKLVAAASQAALGL 609

RESULT 27

ABU57252
ID ABU57252 standard; protein; 609 AA.

XX AC ABU57252;

XX 25-APR-2003 (first entry)

DT 25-APR-2003 (first entry)

XX Human serum albumin protein #1.

XX Human; enzyme; transgenic; drug metabolism; behaviour; mouse;

XX pharmacokinetic assay; pharmacodynamic assay; toxicology; serum albumin;

XX alpha-acidic glycoprotein; CYP; multidrug resistance protein; MRP;

XX uridine diphosphoglucuronosyl transferase; UGT; cytochrome P450.

XX Homo sapiens.

XX WO200283897-A1.

XX 24-OCT-2002.

XX 18-APR-2002; 2002WO-AU000485.

XX 18-APR-2001; 2001AU-00004467.

XX (GENE-) GENE STREAM PTY LTD.

XX Daly JM;

XX WPI; 2003-093021/08.

XX N-PSDB; ABX77157.

XX New transgenic non-human animal expressing a foreign polypeptide

XX associated with drug behavior and/or metabolism, useful for studying the

XX behavior and/or metabolism of a drug in other animals.

XX Disclosure; Page 109-112; 408pp; English.

XX This invention relates to a transgenic non-human animal which may be used

XX for assessing the behaviour and/or metabolism of a drug in another animal

XX and which expresses a foreign polypeptide associated with drug behaviour

XX and/or metabolism. The invention also comprises a nucleic acid construct

XX for use in producing the above transgenic non-human animal and a method

XX of assessing the metabolism and/or behavior of a drug in an animal of

XX interest, comprising administering a test agent to the transgenic animal

XX and conducting analytical tests to determine drug metabolism and/or

XX behaviour. The transgenic animal is useful in studying drug metabolism

CC and/or behavior in other animals. The nucleic acid construct is useful
 CC in producing the above transgenic animal and the methods are used for
 CC producing, breeding and using transgenic animals for pharmacological
 CC (e.g. pharmacokinetic or pharmacodynamic assays) and/or toxicological
 CC studies. Nucleic acid sequences used within the invention are serum
 CC albumin; alpha-acidic glycoprotein; cytochrome P450 (CYP); uridine
 CC diphosphoglucuronosyl transferase (UGT); multidrug resistance proteins
 CC and (MRP's). The present sequence represents a protein sequence used to
 CC create a transgenic animal within the scope of the invention
 XX
 SQ Sequence 609 AA;
 Query Match 100.0%; Score 585; DB 6; Length 609;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLOQCPFDHVKLVNEVTEFAKTCVADSEAE 60
 DB 25 DAHKEVAHRFKDLGEENFKALVLIAPAYLOQCPFDHVKLVNEVTEFAKTCVADSEAE 84
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOKDDNPNLPLVRPEV 120
 DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOKDDNPNLPLVRPEV 144
 QY 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFAPPELLFFAKRYKAAPTECCQADKAACLLP 180
 DB 145 DVNCTAFHDNEETFLKKYLYEIAARRHPYFAPPELLFFAKRYKAAPTECCQADKAACLLP 204
 QY 181 KLDELDEGKASSAKORLKCASLQKFGERAFAKAVARLSORFPKAEVSKLVTDLTK 240
 DB 205 KLDELDEGKASSAKORLKCASLQKFGERAFAKAVARLSORFPKAEVSKLVTDLTK 264
 QY 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKKECEKPLLEKSHCIAEVENDEMPA 300
 DB 265 VHTTECHGDLLECCADRADLAKYICENQDSISSKKECEKPLLEKSHCIAEVENDEMPA 324
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVLLLLAKTYETTLK 360
 DB 325 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVLLLLAKTYETTLK 384
 QY 361 CAAADPHECVAKVDFEFKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420
 DB 385 CAAADPHECVAKVDFEFKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 444
 QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVLNOLCVLHEKTPVSDRVTKCCTES 480
 DB 445 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVLNOLCVLHEKTPVSDRVTKCCTES 504
 QY 481 LVNRRPCFSALEVDVTPKFEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540
 DB 505 LVNRRPCFSALEVDVTPKFEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 564
 QY 541 KEQLKAYMDDFAAFVEKCKKADDDKETCFABEGKKLVAASQAALGL 585
 DB 565 KEQLKAYMDDFAAFVEKCKKADDDKETCFABEGKKLVAASQAALGL 609
 RESULT 28
 ABUS7253
 ID ABUS7253 standard; protein; 609 AA.
 AC
 XX
 AC ABUS7253;
 XX
 DT 25-APR-2003 (first entry)
 XX
 DE Human serum albumin protein #2.
 XX
 KW Human; enzyme; transgenic; drug metabolism; behaviour; mouse;
 KW pharmacokinetic assay; pharmacodynamic assay; toxicology; serum albumin;
 KW alpha-acidic glycoprotein; CYP; multidrug resistance protein; MRP;
 KW uridine diphosphoglucuronosyl transferase; UGT; cytochrome P450.
 XX
 OS Homo sapiens.

XX WO200283897-A1.
 PN 24-OCT-2002.
 XX 18-APR-2002; 2002WO-AU0000485.
 XX 18-APR-2001; 2001AU-00004467.
 XX (GENE-) GENE STREAM PTY LTD.
 PA Dally JM;
 PI WPI; 2003-093021/08.
 DR N-PSDB; ABX77158.
 XX
 PT New transgenic non-human animal expressing a foreign polypeptide
 PT associated with drug behavior and/or metabolism, useful for studying the
 PT behavior and/or metabolism of a drug in other animals.
 XX
 PS Disclosure; Page 115-117; 408pp; English.
 XX
 CC This invention relates to a transgenic non-human animal which may be used
 CC for assessing the behaviour and/or metabolism of a drug in another animal
 CC and which expresses a foreign polypeptide associated with drug behaviour
 CC and/or metabolism. The invention also comprises a nucleic acid construct
 CC for use in producing the above transgenic non-human animal and a method
 CC of assessing the metabolism and/or behavior of a drug in an animal of
 CC interest, comprising administering a test agent to the transgenic animal
 CC and conducting analytical tests to determine drug metabolism and/or
 CC behaviour. The transgenic animal is useful in studying drug metabolism
 CC and/or behaviour in other animals. The nucleic acid construct is useful
 CC in producing the above transgenic animal and the methods are used for
 CC producing, breeding and using transgenic animals for pharmacological
 CC (e.g. pharmacokinetic or pharmacodynamic assays) and/or toxicological
 CC studies. Nucleic acid sequences used within the invention are serum
 CC albumin; alpha-acidic glycoprotein; cytochrome P450 (CYP); uridine
 CC diphosphoglucuronosyl transferase (UGT); multidrug resistance proteins
 CC and (MRP's). The present sequence represents a protein sequence used to
 CC create a transgenic animal within the scope of the invention
 XX
 SQ Sequence 609 AA;
 Query Match 100.0%; Score 585; DB 6; Length 609;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLOQCPFDHVKLVNEVTEFAKTCVADSEAE 60
 DB 25 DAHKEVAHRFKDLGEENFKALVLIAPAYLOQCPFDHVKLVNEVTEFAKTCVADSEAE 84
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOKDDNPNLPLVRPEV 120
 DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOKDDNPNLPLVRPEV 144
 QY 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFAPPELLFFAKRYKAAPTECCQADKAACLLP 180
 DB 145 DVNCTAFHDNEETFLKKYLYEIAARRHPYFAPPELLFFAKRYKAAPTECCQADKAACLLP 204
 QY 181 KLDELDEGKASSAKORLKCASLQKFGERAFAKAVARLSORFPKAEVSKLVTDLTK 240
 DB 205 KLDELDEGKASSAKORLKCASLQKFGERAFAKAVARLSORFPKAEVSKLVTDLTK 264
 QY 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKKECEKPLLEKSHCIAEVENDEMPA 300
 DB 265 VHTTECHGDLLECCADRADLAKYICENQDSISSKKECEKPLLEKSHCIAEVENDEMPA 324
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVLLLLAKTYETTLK 360
 DB 325 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVLLLLAKTYETTLK 384
 QY 361 CAAADPHECVAKVDFEFKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420

385 CAAADPHCYAKVDFEFKPLVEEPQNLKQNCSELFQELGEYKFNQALLVRYTKVQVST 444
421 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVNLQNLVLEKTPVSDRVTKCTES 480
445 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVNLQNLVLEKTPVSDRVTKCTES 504
481 LVNRRPCFSALVDETVYVPKFNATFTFHADICTLSEKQIKKQALVELVVKHFKPAT 540
505 LVNRRPCFSALVDETVYVPKFNATFTFHADICTLSEKQIKKQALVELVVKHFKPAT 564
541 KEQLKAVMDDFAAVFEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
565 KEQLKAVMDDFAAVFEKCKCKADDDKTCFAEKGKLVAAASQAALGL 609

RESULT 29
ADA24217
ID ADA24217 standard; protein; 609 AA.
XX
AC ADA24217;
DT 20-NOV-2003 (first entry)
XX
DE Human serum albumin protein SEQ ID NO:23.
XX
KW therapeutic oligonucleotide; double-stranded RNA; dsRNA; mobile protein;
KW cytosolic; immunosuppressive; virucide; anti-HIV; antibacterial;
KW cardiant; hyperproliferation; cancer; haematological; metastatic;
KW autoimmune disease; infection; endocrine; neural; cardiovascular;
KW pulmonary; reproductive system disorder; endocytosis; metabolic process;
KW human; human serum albumin; HSA.
XX
OS Homo sapiens.
XX
PN WO2003069306-A2.
XX
PD 21-AUG-2003.
XX
PF 13-FEB-2003; 2003WO-US004323.
XX
PR 13-FEB-2002; 2002US-0356053P.
XX
XX (MEDB-) MEDBRIDGE INC.
XX
PA Xie D;
XX
PI WPI; 2003-646491/61.
XX
DR Treating diseases with oligonucleotides or interfering RNA, useful e.g.
PT for cancer or autoimmune diseases, covalently coupled to mobile proteins,
PT in vivo or in vitro.
XX
PS Claim 28; Page 19; 42pp; English.

The present invention describes a method for treating a disease by administering: (a) a therapeutic oligonucleotide (TON) or double-stranded RNA (dsRNA) that includes a reactive group (RG) that can react with a mobile protein (MP) to form a covalent conjugate of TON/dsRNA and MP; or (b) TON or dsRNA already conjugated to MP through a covalent bond. Also described: (i) TON of 15-30 bases that includes (i) a part that binds to target RNA or DNA and (ii) RG; (2) TON of 15-30 bases that includes a part that binds to target RNA or DNA and is conjugated to MP through a covalent link; (3) dsRNA that includes RG; and (4) dsRNA that is conjugated to MP through a covalent link. TON have cytostatic, immunosuppressive, virucide, anti-HIV, antibacterial and cardiac activities. The method is used to treat, or prevent, hyperproliferation of (particularly) cancers, solid or haematological, including prevention of metastatic spread; autoimmune diseases; viral or bacterial infections; endocrine, neural, cardiovascular, pulmonary or reproductive system disorders. Also where TON or dsRNA are labelled, they can be used for diagnosis and monitoring of therapy. When linked to a mobile protein, TON/dsRNA have better cell entry (via endocytosis or other parts of the mobile protein metabolic process) and longer therapeutic life, increased

CC from hours to weeks (the result of increased resistance to nuclease),
CC without loss of affinity for the target. In many cases immune response to
CC TON/dsRNA is also reduced, as is non-specific binding to endogenous
CC proteins. The present sequence represents a human serum albumin protein
CC amino acid sequence which is given in the exemplification of the present
CC invention.

XX
SQ Sequence 609 AA;
Query Match 100.0%; Score 585; DB 7; Length 609;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHREKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEPAKTCVADESAB 60
DB 25 DAHSEVAHREKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEPAKTCVADESAB 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAKAACLLP 180
DB 145 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAKAACLLP 204
QY 181 KLDELREDEGKASSAKQRLKCAASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 240
DB 205 KLDELREDEGKASSAKQRLKCAASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 264
QY 241 VITECHGDLLECADDDRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
DB 265 VITECHGDLLECADDDRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKVNYABAKDVFGLMFLEYARRHPDYSVVLLRLAKTYETTTLEK 360
DB 325 DLPSLAADFVESKDVCKVNYABAKDVFGLMFLEYARRHPDYSVVLLRLAKTYETTTLEK 384
QY 361 CAAADPHCYAKVDFEFKPLVEEPQNLKQNCSELFQELGEYKFNQALLVRYTKVQVST 420
DB 385 CAAADPHCYAKVDFEFKPLVEEPQNLKQNCSELFQELGEYKFNQALLVRYTKVQVST 444
QY 421 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVNLQNLVLEKTPVSDRVTKCTES 480
DB 445 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVNLQNLVLEKTPVSDRVTKCTES 504
QY 481 LVNRRPCFSALVDETVYVPKFNATFTFHADICTLSEKQIKKQALVELVVKHFKPAT 540
DB 505 LVNRRPCFSALVDETVYVPKFNATFTFHADICTLSEKQIKKQALVELVVKHFKPAT 564
QY 541 KEQLKAVMDDFAAVFEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
DB 565 KEQLKAVMDDFAAVFEKCKCKADDDKTCFAEKGKLVAAASQAALGL 609

RESULT 30
ADD06471
ID ADD06471 standard; protein; 609 AA.
XX
AC ADD06471;
DT 01-JAN-2004 (first entry)
XX
DE Human serum albumin protein SEQ ID NO:7.
XX
KW human; chemokine beta1; Ckbl; anti-HIV; neuroprotective; antithyroid;
KW antiarthritic; antirheumatic; immunosuppressive; nootropic;
KW antiinflammatory; antiasclerotic; antiallergic; osteopathic;
KW nephrotrophic; tuberculostatic; virucide; antiatherosclerotic;
KW antimicrobial; infection; HIV; immune disorder; haematopoietic disorder;
KW autoimmune disorder; multiple sclerosis; Grave's disease; arthritis;
KW rheumatoid arthritis; transplant rejection; neurodegenerative disorder;
KW Alzheimer's disease; inflammatory disease; asthma; allergic disorder;
KW inflammatory bowel disease; osteoarthritis; colitis;

KW inflammatory kidney disease; glomerulonephritis; infectious disease;
KW tuberculosis; hepatitis infection; herpes viral infection;
KW viral infection; proliferative disorder; atherosclerosis;
KW human serum albumin; HSA.
XX
OS Homo sapiens.
XX
XX WO200297038-A2.
XX
XX
XX 05-DEC-2002.
XX
XX 24-MAY-2002; 2002WO-US016525.
XX
XX 25-MAY-2001; 2001US-0293212P.
XX
XX (HUYA-) HUMAN GENOME SCI INC.
XX
XX Bell A, Ruben SM;
XX
XX WPI; 2003-140456/13.
XX
XX Novel human chemokine betal protein comprising deletion in amino acids
PT from amino and/or carboxy terminus, and is a fusion protein further
PT comprising human serum albumin, is useful for treating multiple
PT sclerosis, asthma.
XX
XX Disclosure; SEQ ID NO 7; 423pp; English.
XX
XX The present invention describes a human chemokine betal (Ckb1) protein
CC (I) comprising a deletion in amino acid residues from the amino terminus
CC and/or carboxy terminus of the 93 residue amino acid sequence (SI, see
CC ADD06466). (I) has anti-HIV, neuroprotective, antithyroid, antiarthritic,
CC antirheumatic, immunosuppressive, nontropic, antiinflammatory,
CC antiallergic, osteopathic, nephrotrophic, tuberculostatic,
CC virucide, antiatherosclerotic and antimicrobial activities. (I) is useful
CC for preventing infection, preferably viral (human immunodeficiency virus
CC (HIV)) infection, in a cell, by contacting the cell with (I). (I) is also
CC useful for treating a disease, such as HIV infection or immune disorders,
CC haematopoietic disorders, autoimmune disorders, multiple sclerosis,
CC Grave's disease, arthritis, rheumatoid arthritis, transplant rejection,
CC neurodegenerative disorders, Alzheimer's disease, inflammatory disease,
CC asthma, allergic disorders, inflammatory bowel disease, osteoarthritis,
CC colitis, inflammatory kidney diseases, glomerulonephritis, infectious
CC disease, tuberculosis, hepatitis infections, herpes viral infection,
CC viral infection, proliferative disorders or atherosclerosis, in an
CC individual. (I) inhibits or abolishes the ability of HIV to bind to,
CC enter into/fuse with (infect), and/or replicate in CCR5 expressing cells.
CC (I) also acts a CCR5 agonists or antagonists, stimulate chemotaxis of
CC CCR5-expressing cells, inhibit CCR5 ligand binding to a CCR5 molecule, or
CC upregulate or downregulate CCR5 expression. (I) is useful as an
CC immunological probe for the differential identification of the tissues or
CC cell-types. (I)-human serum albumin (HSA) fusion proteins are useful for
CC diagnosing, treating and preventing various disorders in mammals,
CC preferably in humans. (I)-HSA fusion proteins are also useful as
CC molecular weight markers on sodium dodecyl sulfate polyacrylamide gel
CC electrophoresis techniques, for raising antibodies, and to test the
CC biological activities of the Ckb1 protein. (I)-HSA fusion proteins are
CC useful for screening for molecules that bind to the Ckb1 protein portion
CC of the fusion protein. The present sequence is used in the
XX exemplification of the present invention.
XX
XX Sequence 609 AA;
SQ

Query Match 100.0%; Score 585; DB 7; Length 609;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRFKDGLGENFKALVIAFAQYLOQCFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHKSEVAHRFKDGLGENFKALVIAFAQYLOQCFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNERNECFLOHKDDNPNLPLVRPEV 120

85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNERNECFLOHKDDNPNLPLVRPEV 144
121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFPKRYKAAFTCCCOADKAACLLP 180
145 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFPKRYKAAFTCCCOADKAACLLP 204
181 KLDELDEGKASSAKQRLKCAQLQKFGERAFKAWAVARLSQRFPKAEFAVSKLVTDLTK 240
205 KLDELDEGKASSAKQRLKCAQLQKFGERAFKAWAVARLSQRFPKAEFAVSKLVTDLTK 264
241 VHTCCCHGDLLECCADDRADLAKYICENODSISKLKCECKPLEKSHCHCIAEVENDEMPA 300
265 VHTCCCHGDLLECCADDRADLAKYICENQDSISKLKCECKPLEKSHCHCIAEVENDEMPA 324
301 DLPSLAADPFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTLEKC 360
325 DLPSLAADPFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTLEKC 384
361 CAADAPHECHYAKVDFEPKPLVEPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
385 CAADAPHECHYAKVDFEPKPLVEPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 444
421 PTLVEVSRNLGKVGSKCKHPEAKRMPCADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
445 PTLVEVSRNLGKVGSKCKHPEAKRMPCADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
481 LVNRRPCEFALEVDVETVPKEFNATFTFHADICTLSEKERQIKKQTALVELYKHKPKAT 540
505 LVNRRPCEFALEVDVETVPKEFNATFTFHADICTLSEKERQIKKQTALVELYKHKPKAT 564
541 KEOLKAVNDDFAAFVEKCKCKADDKETCFABEGKGLVAASQAALGL 585
565 KEOLKAVNDDFAAFVEKCKCKADDKETCFABEGKGLVAASQAALGL 609
RESULT 31
ADE77205
ID ADE77205 standard; protein; 609 AA.
XX ADE77205;
XX
XX 29-JAN-2004 (first entry)
XX Human protein expressed in a liver disorder #91.
XX human; liver disorder; hyperlipidaemia; hypertension; type II diabetes;
KW tumour; liver; inflammatory disorder; immune response disorder;
KW high-throughput screening; differential gene expression; gene therapy.
XX Homo sapiens.
XX US2003108871-A1.
XX 12-JUN-2003.
XX 30-JUL-2001; 2001US-00919039.
XX 28-JUL-2000; 2000US-0222113P.
XX (KASE/) KASER M R.
XX Kaser MR;
XX WPI; 2004-031227/03.
XX N-PSDB; ADE77204.
XX Composition comprising several cDNAs that are differentially expressed in
PT treated human C3A liver cell cultures, useful for treating liver
PT disorders.
XX Claim 1; SEQ ID NO 370; 41pp; English.
XX The invention relates to a composition comprising several cDNAs that are
CC

CC differentially expressed in a liver disorder. The composition is useful
CC for treating liver disorder such as hyperlipidaemia, hypertension, type
CC II diabetes, tumours of the liver and disorders of the inflammatory and
CC immune response. The composition is useful for a high-throughput method
CC of screening several molecules or compounds to identify a ligand which
CC specifically binds a cDNA. A protein encoded by the cDNA is useful for a
CC high-throughput method for using a protein to screen several molecules or
CC compounds to identify at least one ligand which specifically binds the
CC protein which involves combining the protein encoded by the cDNA with
CC several of molecules or compounds under conditions to allow specific
CC binding, and detecting specific binding between the protein and a
CC molecule or compound, therefore identifying a ligand which specifically
CC binds the protein. The composition is useful for detecting and
CC quantifying differential gene expression, can be used in gene therapy, to
CC formulate prognosis and to design a treatment regimen and to monitor the
CC efficacy of treatment. The present sequence represents the amino acid
CC sequence of a protein encoded by a cDNA differentially expressed in a
CC liver disorder.
XX
XX
SQ Sequence 609 AA;

Query Match 100.0%; Score 585; DB 8; Length 609;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLOQCFFEDHVKLVNEVTFEFAKTCVADESAAE 60
Db 25 DAHSEVAHRFKDLGEENFKALVLIAPQYLOQCFFEDHVKLVNEVTFEFAKTCVADESAAE 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLRLVRPEV 144

QY 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAETECQADKAACLLP 180
Db 145 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAETECQADKAACLLP 204

QY 181 KLDELDEGKASSAKQRLKCSAQKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
Db 205 KLDELDEGKASSAKQRLKCSAQKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 264

QY 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLKECKEPLEKSHCIAEVENDEMPA 300
Db 265 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLKECKEPLEKSHCIAEVENDEMPA 324

QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLVEYARRHPDYSVLLILLAKTYETTLK 360
Db 325 DLPSLAADFVESKDVCKNVAEAKDVLGMFLVEYARRHPDYSVLLILLAKTYETTLK 384

QY 361 CAADPHECAKVPDEFKPLVEEPQMLIKONCELPQLGEYKFNALLVRYTKVPQVST 420
Db 385 CAADPHECAKVPDEFKPLVEEPQMLIKONCELPQLGEYKFNALLVRYTKVPQVST 444

QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504

QY 481 LVNRRPFCFSALEVDVETVPKFEFNAETFTPHADICTLSEKERQIKQTALVELVKHKPKAT 540
Db 505 LVNRRPFCFSALEVDVETVPKFEFNAETFTPHADICTLSEKERQIKQTALVELVKHKPKAT 564

QY 541 KEQLKAWMDDFAFVKECKKADDKETCFAEKGKLVAAASQAALGL 585
Db 565 KEQLKAWMDDFAFVKECKKADDKETCFAEKGKLVAAASQAALGL 609

RESULT 32
AAR39510
ID AAR39510 standard; protein; 610 AA.
XX
XX AAR39510;
AC
XX
DT 25-MAR-2003 (revised)

DT 02-FEB-1994 (first entry)
XX Chimeric human serum albumin.
DE
XX Chimeric protein; albumin; fusion protein; therapeutic polypeptide.
KW
XX Synthetic.
OS
XX
FH Key Location/Qualifiers
FT Misc_diff 610 /note= "Therapeutically active peptide attached to human serum albumin here."
FT
XX
XX W09315199-Al.
XX
XX 05-AUG-1993.
XX
XX 28-JAN-1993; 93WO-FR000085.
XX
XX 31-JAN-1992; 92FR-00001064.
XX (RHON) RHONE POULENC RORER SA.
XX
XX Fleer R, Fournier A, Guitton J, Jung G, Yeh P;
PI WPI; 1993-258677/32.
DR N-PSDB; AAQ46007.
DR
XX
XX New therapeutically active fusion proteins - comprising active polypeptide linked to albumin (variant).
PT
XX
XX Claim 1; Fig 2; 60pp; French.
XX
CC The recombinant polypeptides are plasma-stable versions of the original
CC therapeutically active polypeptide and may be used for the same
CC therapeutic purposes. They may also have enhanced activity and/or
CC reduced side effects. The therapeutic polypeptides used are selected from
CC enzymes, enzyme inhibitors, antigens, antibodies, hormones, blood
CC clotting factors, osteogenesis and/or bone resorption factors,
CC interferons, cytokines, bactericidal or antifungal factors etc. The
CC active portion comprises the whole peptide, a fragment, or a mutant
CC retaining the therapeutic activity. The active portion is coupled to the
CC N- or C-terminus of the albumin. (Updated on 25-MAR-2003 to correct PN
CC field.)
XX
SQ Sequence 610 AA;

Query Match 100.0%; Score 585; DB 2; Length 610;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLOQCFFEDHVKLVNEVTFEFAKTCVADESAAE 60
Db 25 DAHSEVAHRFKDLGEENFKALVLIAPQYLOQCFFEDHVKLVNEVTFEFAKTCVADESAAE 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLRLVRPEV 144

QY 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAETECQADKAACLLP 180
Db 145 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAETECQADKAACLLP 204

QY 181 KLDELDEGKASSAKQRLKCSAQKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
Db 205 KLDELDEGKASSAKQRLKCSAQKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 264

QY 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLKECKEPLEKSHCIAEVENDEMPA 300
Db 265 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLKECKEPLEKSHCIAEVENDEMPA 324

QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLVEYARRHPDYSVLLILLAKTYETTLK 360

[illegible]

PT Novel heterologous fusion protein, useful for treating non-insulin
PT dependent diabetes mellitus or obesity, comprises a glucagon-like peptide
PT 1 compound fused to human albumin or to the Fc portion of an
XX immunoglobulin.
XX
XX Example 6; Page 81; 200pp; English.
XX
XX The invention relates to a heterologous fusion protein comprising a first
CC polypeptide fused to a second polypeptide, where the polypeptides has a N
CC -terminus and a C-terminus and the first polypeptide is a glucagon-like
CC peptide 1 (GLP-1) compound and the second is a human albumin or its
CC analogue or fragment, or the Fc portion of an immunoglobulin (Ig) or its
CC analogue or fragment, where the C-terminus of first polypeptide is fused
CC to the N-terminus of the second polypeptide. The invention is useful for
CC normalising blood glucose levels in mammal, for treating a patient with
CC non-insulin diabetes mellitus or obesity, or for the manufacture of
CC medicament for treating the above mentioned diseases. The present
CC sequence is human serum albumin protein (HSA)
XX
XX Sequence 624 AA;
SQ
Query Match 100.0%; Score 585; DB 6; Length 624;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DAHSEVAHRFKDLGEBNFKALVLI AFAQYLOCCPFEDHVKLVNEVTFKTCVADSAE 60
Db 40 DAHSEVAHRFKDLGEBNFKALVLI AFAQYLOCCPFEDHVKLVNEVTFKTCVADSAE 99
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDDNPNLPLVRREV 120
Db 100 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDDNPNLPLVRREV 159
Qy 121 DVNCTAFHDNEETFLKKYLIEARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 160 DVNCTAFHDNEETFLKKYLIEARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 219
Qy 181 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSQRPFAEYVSKLVTDLTK 240
Db 220 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSQRPFAEYVSKLVTDLTK 279
Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAFVENDEMPA 300
Db 280 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAFVENDEMPA 339
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIARRHPDYSVLLLRKAKTYETTLK 360
Db 340 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIARRHPDYSVLLLRKAKTYETTLK 399
Qy 361 CAADDPHECYAKVDFEFKPLVEBPQNLKONCELFQOLGEYKFQNALVRYTKVPQVST 420
Db 400 CAADDPHECYAKVDFEFKPLVEBPQNLKONCELFQOLGEYKFQNALVRYTKVPQVST 459
Qy 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 460 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 519
Qy 481 LVNRRPCFSALVEVDYVPKFEFNAETFTPHADICTLSEKERQIKKQATLVELVGHKPKAT 540
Db 520 LVNRRPCFSALVEVDYVPKFEFNAETFTPHADICTLSEKERQIKKQATLVELVGHKPKAT 579
Qy 541 KEQLKAVMDDPFAAFVEKCKKADKTCFASGEGKLVAAASQAALGL 585
Db 580 KEQLKAVMDDPFAAFVEKCKKADKTCFASGEGKLVAAASQAALGL 624
RESULT 35
AAE30917
ID AAE30917 standard; protein; 631 AA.
XX
XX AAE30917;
AC
XX
XX
DT 24-FEB-2003 (first entry)

XX Val8-GLP-1-linker-human serum albumin (HSA) fusion protein.
DE Human; glucagon-like peptide 1; GLP-1; albumin; immunoglobulin; Ig;
XX therapy; non-insulin diabetes mellitus; obesity; antidiabetic; anorectic;
KW fusion protein.
XX
XX Homo sapiens.
OS Unidentified.
OS Chimeric.
XX WO200246227-A2.
PD 13-JUN-2002.
XX
XX 29-NOV-2001; 2001WO-US043165.
PF
XX 07-DEC-2000; 2000US-0251954P.
PR
XX (ELIL) LILLY & CO ELI.
PA
XX Glaesner W, Micanovic R, Tschang SR;
PI WPI; 2003-018534/01.
XX
XX Novel heterologous fusion protein, useful for treating non-insulin
PT dependent diabetes mellitus or obesity, comprises a glucagon-like peptide
PT 1 compound fused to human albumin or to the Fc portion of an
PT immunoglobulin.
XX
XX Example 6; Page 80; 200pp; English.
XX
XX The invention relates to a heterologous fusion protein comprising a first
CC polypeptide fused to a second polypeptide, where the polypeptides has a N
CC -terminus and a C-terminus and the first polypeptide is a glucagon-like
CC peptide 1 (GLP-1) compound and the second is a human albumin or its
CC analogue or fragment, or the Fc portion of an immunoglobulin (Ig) or its
CC analogue or fragment, where the C-terminus of first polypeptide is fused
CC to the N-terminus of the second polypeptide. The invention is useful for
CC normalising blood glucose levels in mammal, for treating a patient with
CC non-insulin diabetes mellitus or obesity, or for the manufacture of
CC medicament for treating the above mentioned diseases. The present
CC sequence is GLP-1 fusion protein
XX
XX Sequence 631 AA;
SQ
Query Match 100.0%; Score 585; DB 6; Length 631;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DAHSEVAHRFKDLGEBNFKALVLI AFAQYLOCCPFEDHVKLVNEVTFKTCVADSAE 60
Db 47 DAHSEVAHRFKDLGEBNFKALVLI AFAQYLOCCPFEDHVKLVNEVTFKTCVADSAE 106
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDDNPNLPLVRREV 120
Db 107 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDDNPNLPLVRREV 166
Qy 121 DVNCTAFHDNEETFLKKYLIEARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 167 DVNCTAFHDNEETFLKKYLIEARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 226
Qy 181 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSQRPFAEYVSKLVTDLTK 240
Db 227 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSQRPFAEYVSKLVTDLTK 286
Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAFVENDEMPA 300
Db 287 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAFVENDEMPA 346
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIARRHPDYSVLLLRKAKTYETTLK 360
Db 347 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIARRHPDYSVLLLRKAKTYETTLK 406

	Matches	585;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;		
QY	1	DAHSEVAHRRFKD	GEENFKALVLI	FAQVLO	QCPEDHVKLV	NEVTEFAKTCVA	DESAAE	60				
Db	56	DAHSEVAHRRFKD	GEENFKALVLI	FAQVLO	QCPEDHVKLV	NEVTEFAKTCVA	DESAAE	115				
QY	61	NCDKSLHTL	FGDKLCTVAT	RETYGE	MADCCAKQ	EPERNECF	LQHKDDNPNL	RLVRPEV	120			
Db	116	NCDKSLHTL	FGDKLCTVAT	RETYGE	MADCCAKQ	EPERNECF	LQHKDDNPNL	RLVRPEV	175			
QY	121	DVMCTAFHND	BEETLKKYL	YEIAARRH	PFYAP	PELLFP	AKRYKAAFT	TECCOAAKA	CLLP	180		
Db	176	DVMCTAFHND	BEETLKKYL	YEIAARRH	PFYAP	PELLFP	AKRYKAAFT	TECCOAAKA	CLLP	235		
QY	181	KLDELRLDEG	KASSAKQRL	KCASLQ	KQGERAF	KAWAVARL	SQRPKAE	FAEVS	KLVDLT	240		
Db	236	KLDELRLDEG	KASSAKQRL	KCASLQ	KQGERAF	KAWAVARL	SQRPKAE	FAEVS	KLVDLT	295		
QY	241	VHTECCGDL	LLECAD	DRADLA	KYICENQ	DSISSK	LECCBK	PLLEKSH	CHIAE	VENDEMPA	300	
Db	296	VHTECCGDL	LLECAD	DRADLA	KYICENQ	DSISSK	LECCBK	PLLEKSH	CHIAE	VENDEMPA	355	
QY	301	DLPSLAAD	FVESKDV	CKNYAE	AKDVFL	GMFLYE	YARRHP	DYSV	VLLLR	LAKTYETTLEK	360	
Db	356	DLPSLAAD	FVESKDV	CKNYAE	AKDVFL	GMFLYE	YARRHP	DYSV	VLLLR	LAKTYETTLEK	415	
QY	361	CAAADPHE	CYAKVF	DEPKPL	VEEPQ	NLIKON	CELF	FEQ	LGHYK	QFONALLV	RYTKVPQVST	420
Db	416	CAAADPHE	CYAKVF	DEPKPL	VEEPQ	NLIKON	CELF	FEQ	LGHYK	QFONALLV	RYTKVPQVST	475
QY	421	PTLVEVSR	NLKVGS	CKCKH	PEAKR	MPCAED	YLSV	LNQL	CVLH	EKTPVSDR	VTKCTES	480
Db	476	PTLVEVSR	NLKVGS	CKCKH	PEAKR	MPCAED	YLSV	LNQL	CVLH	EKTPVSDR	VTKCTES	535
QY	481	LVNRRPCF	SALEVD	ETVVPK	EFNA	ETFTFH	ADICTL	SEKER	QIKKQ	TALVELV	KHKPKAT	540
Db	536	LVNRRPCF	SALEVD	ETVVPK	EFNA	ETFTFH	ADICTL	SEKER	QIKKQ	TALVELV	KHKPKAT	595
QY	541	KEQLKAVM	DDFAAF	VEKCK	ADDK	ETCF	AE	EGK	KLVA	ASQA	ALGL	585
Db	596	KEQLKAVM	DDFAAF	VEKCK	ADDK	ETCF	AE	EGK	KLVA	ASQA	ALGL	640

RESULT 37
AAE30918
ID AAE30918 standard; protein; 640 AA.
XX
AC AAE30918;
XX
DT 24-FEB-2003 (first entry)
XX
DE Gly8-Glu22-GLP-1-CEX-linker-human serum albumin (HSA) fusion protein.
XX
KW Human; glucagon-like peptide 1; GLP-1; albumin; immunoglobulin; Ig;
KW therapy; non-insulin diabetes mellitus; obesity; antidiabetic; anorectic;
KW fusion protein.
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
PN WO200246227-A2.
XX
PD 13-JUN-2002.
XX
PF 29-NOV-2001; 2001WO-US043165.
XX
PR 07-DEC-2000; 2000US-0251954P.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI Glaesner W, Micanovic R, Tschang SR;
XX

QY	361	CAAADPHE	CYAKVF	DEPKPL	VEEPQ	NLIKON	CELF	FEQ	LGHYK	QFONALLV	RYTKVPQVST	420
Db	407	CAAADPHE	CYAKVF	DEPKPL	VEEPQ	NLIKON	CELF	FEQ	LGHYK	QFONALLV	RYTKVPQVST	466
QY	421	PTLVEVSR	NLKVGS	CKCKH	PEAKR	MPCAED	YLSV	LNQL	CVLH	EKTPVSDR	VTKCTES	480
Db	467	PTLVEVSR	NLKVGS	CKCKH	PEAKR	MPCAED	YLSV	LNQL	CVLH	EKTPVSDR	VTKCTES	526
QY	481	LVNRRPCF	SALEVD	ETVVPK	EFNA	ETFTFH	ADICTL	SEKER	QIKKQ	TALVELV	KHKPKAT	540
Db	527	LVNRRPCF	SALEVD	ETVVPK	EFNA	ETFTFH	ADICTL	SEKER	QIKKQ	TALVELV	KHKPKAT	586
QY	541	KEQLKAVM	DDFAAF	VEKCK	ADDK	ETCF	AE	EGK	KLVA	ASQA	ALGL	585
Db	587	KEQLKAVM	DDFAAF	VEKCK	ADDK	ETCF	AE	EGK	KLVA	ASQA	ALGL	631

RESULT 36
AAE30920
ID AAE30920 standard; protein; 640 AA.
XX
AC AAE30920;
XX
DT 24-FEB-2003 (first entry)
XX
DE Exendin-4-linker-human serum albumin (HSA) fusion protein.
XX
KW Human; glucagon-like peptide 1; GLP-1; albumin; immunoglobulin; Ig;
KW therapy; non-insulin diabetes mellitus; obesity; antidiabetic; anorectic;
KW fusion protein.
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
PN WO200246227-A2.
XX
PD 13-JUN-2002.
XX
PF 29-NOV-2001; 2001WO-US043165.
XX
PR 07-DEC-2000; 2000US-0251954P.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI Glaesner W, Micanovic R, Tschang SR;
XX
DR WPI; 2003-018534/01.
XX
PT Novel heterologous fusion protein, useful for treating non-insulin
PT dependent diabetes mellitus or obesity, comprises a glucagon-like peptide
PT 1 compound fused to human albumin or to the Fc portion of an
PT immunoglobulin.
XX
PS Example 6; Page 81-82; 200pp; English.
XX
CC The invention relates to a heterologous fusion protein comprising a first
CC polypeptide fused to a second polypeptide, where the polypeptides has a N
CC -terminus and a C-terminus and the first polypeptide is a glucagon-like
CC peptide 1 (GLP-1) compound and the second is a human albumin or its
CC analogue or fragment, or the Fc portion of an immunoglobulin (Ig) or its
CC to the N-terminus of the second polypeptide. The invention is useful for
CC normalising blood glucose levels in mammal, for treating a patient with
CC non-insulin diabetes mellitus or obesity, or for the manufacture of
CC medicament for treating the above mentioned diseases. The present
CC sequence is a fusion protein of the invention
XX
SQ Sequence 640 AA;
Query Match 100.0%; Score 585; DB 6; Length 640;
Best Local Similarity 100.0%; Pred. No. 0;

DR WPI; 2003-018534/01.

XX Novel heterologous fusion protein, useful for treating non-insulin
PT dependent diabetes mellitus or obesity, comprises a glucagon-like peptide
PT 1 compound fused to human albumin or to the FC portion of an
XX immunoglobulin.

XX Example 6; Page 81; 200pp; English.

XX The invention relates to a heterologous fusion protein comprising a first
CC polypeptide fused to a second polypeptide, where the polypeptides has a N
CC -terminus and a C-terminus and the first polypeptide is a glucagon-like
CC peptide 1 (GLP-1) compound and the second is a human albumin or its
CC analogue or fragment, where the C-terminus of an immunoglobulin (Ig) or its
CC analogue or fragment, where the C-terminus of first polypeptide is fused
CC to the N-terminus of the second polypeptide. The invention is useful for
CC normalising blood glucose levels in mammal. For treating a patient with
CC non-insulin diabetes mellitus or obesity or for the manufacture of
CC medicament for treating the above mentioned diseases. The present
CC sequence is GLP-1 fusion protein

XX Sequence 640 AA;

Query Match 100.0%; Score 585; DB 6; Length 640;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGGENFKALVLIAPAYLOQCPEDHVKLVNVEVTFKTCVADSSAE 60

Db 56 DAHKSEVAHRFKDLGGENFKALVLIAPAYLOQCPEDHVKLVNVEVTFKTCVADSSAE 115

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDNDNPLRLVREPV 120

Db 116 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDNDNPLRLVREPV 175

QY 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYYPAPELJFFAKRYKAAFTCCQAAADKAACLLP 180

Db 176 DVNCTAFHDNEETFLKKYLYEIAARRHPYYPAPELJFFAKRYKAAFTCCQAAADKAACLLP 235

QY 181 KLDELDEGKASAKORLKCASIQKGERAFKAWAVARLSQRFPAEYKSLVTDLTG 240

Db 236 KLDELDEGKASAKORLKCASIQKGERAFKAWAVARLSQRFPAEYKSLVTDLTG 295

QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCKEPRLBKSHCIAEVENDEMPA 300

Db 296 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCKEPRLBKSHCIAEVENDEMPA 355

QY 301 DLPSLAADFVSKDVCNKYAEAKDVFGLMFLEYARRHPDYVSVLLRLAKTYETTLK 360

Db 356 DLPSLAADFVSKDVCNKYAEAKDVFGLMFLEYARRHPDYVSVLLRLAKTYETTLK 415

QY 361 CAADPHCEYAKVDFEFPKPLVEEPQNLKQNCLEFQGLGEYKFQNALVRYTKVPQVST 420

Db 416 CAADPHCEYAKVDFEFPKPLVEEPQNLKQNCLEFQGLGEYKFQNALVRYTKVPQVST 475

QY 421 PTLVEVSRNLGKVGSKCKKHPKAPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

Db 476 PTLVEVSRNLGKVGSKCKKHPKAPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 535

QY 481 LVNRRPFCFSALEVDVTVPKFNAETFTFHADICTLSEKEROIKKQATALVELVKKHKPAT 540

Db 536 LVNRRPFCFSALEVDVTVPKFNAETFTFHADICTLSEKEROIKKQATALVELVKKHKPAT 595

QY 541 KEQKAVMDFAFVPEVCKKADKTCFPAEKGKLVAAQAAIGL 585

Db 596 KEQKAVMDFAFVPEVCKKADKTCFPAEKGKLVAAQAAIGL 640

RESULT 38

ADD06597

ID ADD06597 standard; protein; 651 AA.

XX

AC ADD06597;

01-JAN-2004 (first entry)

Human Ckb1-HAS fusion protein construct secreted protein SEQ ID NO:133.
human; chemokine betal; Ckb1; anti-HIV; neuroprotective; antithyroid;
antiarthritic; antirheumatic; immunosuppressive; neutropic;
antiinflammatory; antiasthmatic; antiallergic; osteopathic;
KW nephrotropic; tuberculostatic; virucide; antiatherosclerotic;
KW antimicrobial; infection; HIV; immune disorder; haematopoietic disorder;
KW autoimmune disorder; multiple sclerosis; Grave's disease; arthritis;
KW rheumatoid arthritis; transplant rejection; neurodegenerative disorder;
KW Alzheimer's disease; inflammatory disease; asthma; allergic disorder;
KW inflammatory bowel disease; osteoarthritis; colitis;
KW inflammatory kidney disease; glomerulonephritis; infectious disease;
KW tuberculosis; hepatitis infection; herpes viral infection;
KW viral infection; proliferative disorder; atherosclerosis;
KW human serum albumin; HSA.

Synthetic.

Homo sapiens.

WO200297038-A2.

05-DEC-2002.

24-MAY-2002; 2002WO-US016525.

25-MAY-2001; 2001US-0293212P.

(HUMA-) HUMAN GENOME SCI INC.

Bell A, Ruben SM;

WPI; 2003-140456/13.

N-PSDB; ADD06601.

Novel human chemokine betal protein comprising deletion in amino acids
from amino and/or carboxy terminus, and is a fusion protein further
comprising human serum albumin, is useful for treating multiple
sclerosis, asthma.

Example 1; SEQ ID NO 133; 423pp; English.

The present invention describes a human chemokine betal (Ckb1) protein
(I) comprising a deletion in amino acid residues from the amino terminus
and/or carboxy terminus of the 93 residue amino acid sequence (S1, see
ADD06466). (I) has anti-HIV, neuroprotective, antithyroid, antiarthritic,
antiasthmatic, immunosuppressive, neutropic, antiinflammatory,
virucide, antiatherosclerotic and antimicrobial activities. (I) is useful
for preventing infection, preferably viral (human immunodeficiency virus
(HIV) infection, in a cell, by contacting the cell with (I). (I) is also
useful for treating a disease, such as HIV infection or immune disorders,
haematopoietic disorders, autoimmune disorders, multiple sclerosis,
Grave's disease, arthritis, rheumatoid arthritis, transplant rejection,
neurodegenerative disorders, Alzheimer's disease, inflammatory disease,
asthma, allergic disorders, inflammatory bowel disease, osteoarthritis,
colitis, inflammatory kidney diseases, glomerulonephritis, infectious
disease, tuberculosis, hepatitis infections, herpes viral infection,
viral infection, proliferative disorders or atherosclerosis, in an
individual. (I) inhibits or abolishes the ability of HIV to bind to,
enter into/fuse with (infect), and/or replicate in CCR5 expressing cells.
(I) also acts a CCR5 agonists or antagonists, stimulate chemotaxis of
CCR5-expressing cells, inhibit CCR5 ligand binding to a CCR5 molecule, or
upregulate or downregulate CCR5 expression. (I) is useful as an
immunological probe for the differential identification of the tissues or
cell-types. (I)-human serum albumin (HSA) fusion proteins are useful for
diagnosing, treating and preventing various disorders in mammals,
preferably in humans. (I)-HSA fusion proteins are also useful as
molecular weight markers on sodium dodecyl sulfate polyacrylamide gel
electrophoresis techniques, for raising antibodies, and to test the
biological activities of the Ckb1 protein. (I)-HSA fusion proteins are

CC useful for screening for molecules that bind to the Ckbl protein portion
CC of the fusion protein. The present sequence is used in the
CC exemplification of the present invention.

XX
SQ Sequence 651 AA;

Query Match 100.0%; Score 585; DB 7; Length 651;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQQCFEDHVKLVNVEVTEFAKTCVADSEAE 60
DB 67 DAHKEVAHRFKDLGEENFKALVLIAPAYLQQCFEDHVKLVNVEVTEFAKTCVADSEAE 126
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNDPNLPLRVPEV 120
DB 127 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNDPNLPLRVPEV 186
QY 121 DVMCTAFHDNEETFLKKLYEIAIRHPHYFAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 187 DVMCTAFHDNEETFLKKLYEIAIRHPHYFAPPELLFFAKRYKAAFTCCQAADKAAACLLP 246
QY 181 KLDELDEGKASSAKORLKCASLQKFGERAFKANAVARLSORFPAKFAEVSKLVTDLTK 240
DB 247 KLDELDEGKASSAKORLKCASLQKFGERAFKANAVARLSORFPAKFAEVSKLVTDLTK 306
QY 241 VHTCCGDLLECCADRDALAKYICENQDSISSKLKECCEKPLLEKSHCTAEVNDMPA 300
DB 307 VHTCCGDLLECCADRDALAKYICENQDSISSKLKECCEKPLLEKSHCTAEVNDMPA 366
QY 301 DPLSLAADVESKDVCKNVAEKVFLGMLFYEYARRHPDYSVLLRLAKTYETTLLEK 360
DB 367 DPLSLAADVESKDVCKNVAEKVFLGMLFYEYARRHPDYSVLLRLAKTYETTLLEK 426
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLKQNCLEPEQLGEYKFQNALLVYTKVQVST 420
DB 427 CAAADPHECYAKVDFEFKPLVEEPQNLKQNCLEPEQLGEYKFQNALLVYTKVQVST 486
QY 421 PTLVEVSRNLGVSKCCCHPAKMPACABDLSVLNQLCVLHEKTPSVDRVTKCCTES 480
DB 487 PTLVEVSRNLGVSKCCCHPAKMPACABDLSVLNQLCVLHEKTPSVDRVTKCCTES 546
QY 481 LVNRRFCFSALEVDEYVYKPFNAETFFHADICTLSEKERIKQKQALVLELVKHKPKAT 540
DB 547 LVNRRFCFSALEVDEYVYKPFNAETFFHADICTLSEKERIKQKQALVLELVKHKPKAT 606
QY 541 KEOLKAVMDDFAFVKECKCKADDKETCPAEKGKLVAAASQAALGL 585
DB 607 KEOLKAVMDDFAFVKECKCKADDKETCPAEKGKLVAAASQAALGL 651

RESULT 39
ADD06596
ID ADD06596 standard; protein; 652 AA.
XX
AC ADD06596;
XX
XX
DT 01-JAN-2004 (first entry)
XX
XX
DE Human Ckbl-HAS fusion protein construct secreted protein SEQ ID NO:132.

KW human; chemokine betal; Ckbl; anti-HIV; neuroprotective; antithyroid;
KW antiarthritic; antirheumatic; immunosuppressive; neutropic;
KW antiinflammatory; antiasthmatic; antiallergic; osteopathic;
KW nephrotrophic; tuberculostatic; virucide; antiatherosclerotic;
KW antimicrobial; infection; HIV; immune disorder; haematopoietic disorder;
KW autoimmune disorder; multiple sclerosis; Grave's disease; arthritis;
KW rheumatoid arthritis; transplant rejection; neurodegenerative disorder;
KW Alzheimer's disease; inflammatory disease; asthma; allergic disorder;
KW inflammatory bowel disease; osteoarthritis; colitis;
KW inflammatory kidney disease; glomerulonephritis; infection;
KW tuberculosis; hepatitis infection; herpes viral infection;
KW viral infection; proliferative disorder; atherosclerosis;

KW human serum albumin; HSA.
XX Synthetic.
OS Homo sapiens.

XX
PN WO200297038-A2.
XX
PD 05-DEC-2002.
XX
PF 24-MAY-2002; 2002WO-US016525.
XX
PR 25-MAY-2001; 2001US-0293212P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Bell A, Ruben SM;
XX
DR WPI; 2003-140456/13.
DR N-PSDB; ADD06600.

XX Novel human chemokine betal protein comprising deletion in amino acids
PT from amino and/or carboxy terminus, and is a fusion protein further
PT comprising human serum albumin, is useful for treating multiple
PT sclerosis, asthma.

XX Example 1; SEQ ID NO 132; 423pp; English.

XX The present invention describes a human chemokine betal (Ckbl) protein
CC (I) comprising a deletion in amino acid residues from the amino terminus
CC and/or carboxy terminus of the 93 residue amino acid sequence (S1, see
CC ADD06466). (I) has anti-HIV, neuroprotective, antithyroid, antiarthritic,
CC antirheumatic, immunosuppressive, neutropic, antinflammatory,
CC antiasthmatic, antiallergic, osteopathic, nephrotrophic, tuberculostatic,
CC virucide, antiatherosclerotic and antimicrobial activities. (I) is useful
CC for preventing infection, preferably viral (human immunodeficiency virus
CC (HIV)) infection, in a cell, by contacting the cell with (I). (I) is also
CC useful for treating a disease, such as HIV infection or immune disorders,
CC haematopoietic disorders, autoimmune disorders, multiple sclerosis,
CC Grave's disease, arthritis, rheumatoid arthritis, transplant rejection,
CC neurodegenerative disorders, Alzheimer's disease, inflammatory disease,
CC asthma, allergic disorders, inflammatory bowel disease, osteoarthritis,
CC colitis, inflammatory kidney disease, glomerulonephritis, infectious
CC disease, tuberculosis, hepatitis infections, herpes viral infection,
CC viral infection, proliferative disorders or atherosclerosis, in an
CC individual. (I) inhibits or abolishes the ability of HIV to bind to,
CC enter into/fuse with (infect), and/or replicate in CCR5 expressing cells.
CC (I) also acts a CCR5 agonists or antagonists, stimulate chemotaxis of
CC CCR5-expressing cells, inhibit CCR5 ligand binding to a CCR5 molecule, or
CC upregulate or downregulate CCR5 expression. (I) is useful as an
CC immunological probe for the differential identification of the tissues or
CC cell-types. (I)-human serum albumin (HSA) fusion proteins are useful for
CC diagnosing, treating and preventing various disorders in mammals,
CC preferably in humans. (I)-HSA fusion proteins are also useful as
CC molecular weight markers on sodium dodecyl sulfate polyacrylamide gel
CC electrophoresis techniques, for raising antibodies, and to test the
CC biological activities of the Ckbl protein. (I)-HSA fusion proteins are
CC useful for screening for molecules that bind to the Ckbl protein portion
CC of the fusion protein. The present sequence is used in the
CC exemplification of the present invention.

XX
XX Sequence 652 AA;

Query Match 100.0%; Score 585; DB 7; Length 652;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQQCFEDHVKLVNVEVTEFAKTCVADSEAE 60
DB 68 DAHKEVAHRFKDLGEENFKALVLIAPAYLQQCFEDHVKLVNVEVTEFAKTCVADSEAE 127
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNDPNLPLRVPEV 120
DB 128 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNDPNLPLRVPEV 187

QY 121 DVNCTAFHDNEETFLKKYLYEYIARRHPYFAPPELLFFAKRYKAAFTBCCQAADKAAACLLP 180
Db 188 DVNCTAFHDNEETFLKKYLYEYIARRHPYFAPPELLFFAKRYKAAFTBCCQAADKAAACLLP 247
QY 181 KLDELDEGKASSAKORLKCASLQKFGGERAFKAWAVARLSORFFKABFAEYKSLVTLDTLK 240
Db 248 KLDELDEGKASSAKORLKCASLQKFGGERAFKAWAVARLSORFFKABFAEYKSLVTLDTLK 307
QY 241 VHTCCCHGDLLECADDRLADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
Db 308 VHTCCCHGDLLECADDRLADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 367
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360
Db 368 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 427
QY 361 CAAADPHECYAKVDFBFLVEEPQNLKQNCLEFEQLEGEYKFNQALLVRYTKVPQVST 420
Db 428 CAAADPHECYAKVDFBFLVEEPQNLKQNCLEFEQLEGEYKFNQALLVRYTKVPQVST 487
QY 421 PTLVEVSRNLGVSKCCGHPKAPCAEDYLSVVLNQLCVLHEKTPVSRVTKCTTES 480
Db 488 PTLVEVSRNLGVSKCCGHPKAPCAEDYLSVVLNQLCVLHEKTPVSRVTKCTTES 547
QY 481 LVNRRPCFSALEVDETYVPKFNAAETFTPHADICTLSEKERQIKKQTALVELVKEPKAT 540
Db 548 LVNRRPCFSALEVDETYVPKFNAAETFTPHADICTLSEKERQIKKQTALVELVKEPKAT 607
QY 541 KEQLKAWMDFAAFVBEKCKKADDKETCFABEGKLVAAQAALGL 585
Db 608 KEQLKAWMDFAAFVBEKCKKADDKETCFABEGKLVAAQAALGL 652

RESULT 40
ADD06595
ID ADD06595 standard; protein; 553 AA.
AC ADD06595;
XX
XX
DT 01-JAN-2004 (first entry)
XX
DE Human Ckb1-HAS fusion protein construct secreted protein SEQ ID NO:131.
XX
KW human; chemokine betal; Ckb1; anti-HIV; neuroprotective; antithyroid;
KW antiarthritic; antirheumatic; immunosuppressive; nontropic;
KW antiinflammatory; antiallergic; osteopathic;
KW nephrotrophic; tuberculostatic; virucide; antiatherosclerotic;
KW antimicrobial; infection; HIV; immune disorder; haematopoietic disorder;
KW autoimmune disorder; multiple sclerosis; Grave's disease; arthritis;
KW rheumatoid arthritis; transplant rejection; neurodegenerative disorder;
KW Alzheimer's disease; inflammatory disease; asthma; allergic disorder;
KW inflammatory bowel disease; osteoarthritis; colitis;
KW inflammatory kidney disease; glomerulonephritis; infectious disease;
KW tuberculosis; hepatitis infection; herpes viral infection;
KW viral infection; proliferative disorder; atherosclerosis;
KW human serum albumin; HSA.
XX
OS Synthetic.
OS Homo sapiens.
XX
XX WO200297038-A2.
PN
XX
XX
PD 05-DEC-2002.
XX
XX 24-MAY-2002; 2002WO-US016525.
PF
XX
PR 25-MAY-2001; 2001US-0293212P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Bell A, Ruben SM;
PI
XX

DR WPI: 2003-140456/13.
XX N-PSDB; ADD06599.
PT Novel human chemokine betal protein comprising deletion in amino acids
PT from amino and/or carboxy terminus, and is a fusion protein further
PT comprising human serum albumin, is useful for treating multiple
PT sclerosis, asthma.
XX
XX Example 1; SEQ ID NO 131; 423pp; English.
PS
XX The present invention describes a human chemokine betal (Ckb1) protein
CC (I) comprising a deletion in amino acid residues from the amino terminus
CC and/or carboxy terminus of the 93 residue amino acid sequence (SI, see
CC ADD06466). (I) has anti-HIV, neuroprotective, antithyroid, antiarthritic,
CC antirheumatic, immunosuppressive, nontropic, antiinflammatory,
CC antiallergic, osteopathic, nephrotrophic, tuberculostatic,
CC virucide, antiatherosclerotic and antimicrobial activities. (I) is useful
CC for preventing infection, preferably viral (human immunodeficiency virus
CC (HIV)) infection, in a cell, by contacting the cell with (I). (I) is also
CC useful for treating a disease, such as HIV infection or immune disorders,
CC haematopoietic disorders, autoimmune disorders, multiple sclerosis,
CC Grave's disease, arthritis, Rheumatoid arthritis, transplant rejection,
CC neurodegenerative disorders, Alzheimer's disease, inflammatory disease,
CC asthma, allergic disorders, inflammatory bowel disease, osteoarthritis,
CC colitis, inflammatory kidney diseases, glomerulonephritis, infectious
CC disease, tuberculosis, hepatitis infections, herpes viral infection,
CC viral infection, proliferative disorders or atherosclerosis, in an
CC individual. (I) inhibits or abolishes the ability of HIV to bind to,
CC enter into/fuse with (infect), and/or replicate in CCR5 expressing cells.
CC (I) also acts a CCR5 agonists or antagonists, stimulate chemotaxis of
CC CCR5-expressing cells, inhibit CCR5 ligand binding to a CCR5 molecule, or
CC upregulate or downregulate CCR5 expression. (I) is useful as an
CC immunological probe for the differential identification of the tissues or
CC cell-types. (I)-human serum albumin (HSA) fusion proteins are useful for
CC diagnosing, treating and preventing various disorders in mammals,
CC preferably in humans. (I)-HSA fusion proteins are also useful as
CC molecular weight markers on sodium dodecyl sulfate polyacrylamide gel
CC electrophoresis techniques for raising antibodies, and to test the
CC biological activities of the Ckb1 protein. (I)-HSA fusion proteins are
CC useful for screening for molecules that bind to the Ckb1 protein portion
CC of the fusion protein. The present sequence is used in the
CC exemplification of the present invention.
XX
XX Sequence 653 AA;
QY Query Match 100.0%; Score 585; DB 7; Length 653;
Db Best Local Similarity 100.0%; Pred. No. 0;
QY Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 DAKHSEVAHRFKDLGEENFKALVLIAPFAQYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 60
Db 69 DAKHSEVAHRFKDLGEENFKALVLIAPFAQYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 128
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVRPEV 120
Db 129 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPRLVRPEV 188
QY 121 DVNCTAFHDNEETFLKKYLYEYIARRHPYFAPPELLFFAKRYKAAFTBCCQAADKAAACLLP 180
Db 189 DVNCTAFHDNEETFLKKYLYEYIARRHPYFAPPELLFFAKRYKAAFTBCCQAADKAAACLLP 248
QY 181 KLDELDEGKASSAKORLKCASLQKFGGERAFKAWAVARLSORFFKABFAEYKSLVTLDTLK 240
Db 249 KLDELDEGKASSAKORLKCASLQKFGGERAFKAWAVARLSORFFKABFAEYKSLVTLDTLK 308
QY 241 VHTCCCHGDLLECADDRLADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
Db 309 VHTCCCHGDLLECADDRLADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 368
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360
Db 369 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 428

QY 361 CAADPHCEYAKVDFEFKPLVEEPONLIKONCELFQELGEYKFQNALLVRYTKVPQVST 420
DB 429 CAADPHCEYAKVDFEFKPLVEEPONLIKONCELFQELGEYKFQNALLVRYTKVPQVST 488
QY 421 PTLVSVSNLGVSKGCKCKHPEAKMPCAEYDLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 489 PTLVSVSNLGVSKGCKCKHPEAKMPCAEYDLSVLNQLCVLHEKTPVSDRVTKCCTES 548
QY 481 LVNRPFCFSALEVDVTVYPKGFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540
DB 549 LVNRPFCFSALEVDVTVYPKGFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 608
QY 541 KEQLKAVMDDFPAAFEVKCKCKADDDKTCFAEBGKGLVAASQAALGL 585
DB 609 KEQLKAVMDDFPAAFEVKCKCKADDDKTCFAEBGKGLVAASQAALGL 653

RESULT 41

ADD06594

ID ADD06594 standard; protein; 656 AA.

XX

AC ADD06594;

XX

DT 01-JAN-2004 (first entry)

XX

DE Human Ckb1-HAS fusion protein construct secreted protein SEQ ID NO:130.

XX

KW human; chemokine beta1; Ckb1; anti-HIV; neuroprotective; antithyroid;
KW antithyroid; antirheumatic; immunosuppressive; nootropic;
KW antiinflammatory; antiasthmatic; antiallergic; osteopathic;
KW nephrotrophic; tuberculostatic; virucide; antiatherosclerotic;
KW antimicrobial; infection; HIV; immune disorder; haematopoietic disorder;
KW autoimmune disorder; multiple sclerosis; Grave's disease; arthritis;
KW rheumatoid arthritis; transplant rejection; neurodegenerative disorder;
KW Alzheimer's disease; inflammatory disease; asthma; allergic disorder;
KW inflammatory bowel disease; osteoarthritis; colitis;
KW inflammatory kidney disease; glomerulonephritis; infectious disease;
KW tuberculosis; hepatitis infection; herpes viral infection;
KW viral infection; proliferative disorder; atherosclerosis;
KW human serum albumin; HSA.

XX

OS Synthetic.

OS Homo sapiens.

XX

XX WO200297038-A2.

XX

PD 05-DEC-2002.

XX

PF 24-MAY-2002; 2002WO-US016525.

XX

PR 25-MAY-2001; 2001US-0293212P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Bell A, Ruben SM;

XX

XX WPI; 2003-140456/13.

XX

DR N-PSDB; ADD06598.

XX

PT Novel human chemokine beta1 protein comprising deletion in amino acids
PT from amino and/or carboxy terminus, and is a fusion protein further
PT comprising human serum albumin, is useful for treating multiple
PT sclerosis, asthma.

XX

PS Example 1; SEQ ID NO 130; 423pp; English.

XX

CC The present invention describes a human chemokine beta1 (Ckb1) protein
CC (1) comprising a deletion in amino acid residues from the amino terminus
CC and/or carboxy terminus of the 93 residue amino acid sequence (S1, see
CC ADD05466). (I) has anti-HIV, neuroprotective, antithyroid, antirheumatic,
CC antiasthmatic, immunosuppressive, nootropic, antiinflammatory,
CC antiasthmatic, antiallergic, osteopathic, nephrotrophic, tuberculostatic,
CC virucide, antiatherosclerotic and antimicrobial activities. (I) is useful

CC for preventing infection, preferably viral (human immunodeficiency virus
CC (HIV)) infection, in a cell, by contacting the cell with (I). (I) is also
CC useful for treating a disease, such as HIV infection or immune disorders,
CC haematopoietic disorders, autoimmune disorders, multiple sclerosis,
CC Grave's disease, arthritis, rheumatoid arthritis, transplant rejection,
CC neurodegenerative disorders, Alzheimer's disease, inflammatory disease,
CC asthma, allergic disorders, inflammatory bowel disease, osteoarthritis,
CC colitis, inflammatory kidney diseases, glomerulonephritis, infectious
CC disease, tuberculosis, hepatitis infections, herpes viral infection,
CC viral infection, proliferative disorders or atherosclerosis, in an
CC individual. (I) inhibits or abolishes the ability of HIV to bind to,
CC enter into/fuse with (infect), and/or replicate in CCR5 expressing cells.
CC (I) also acts as CCR5 agonists or antagonists, stimulate chemotaxis of
CC CCR5-expressing cells, inhibit CCR5 ligand binding to a CCR5 molecule, or
CC upregulate or downregulate CCR5 expression. (I) is useful as an
CC immunological probe for the differential identification of the tissues or
CC cell-types. (I)-human serum albumin (HSA) fusion proteins are useful for
CC diagnosing, treating and preventing various disorders in mammals,
CC preferably in humans. (I)-HSA fusion proteins are also useful as
CC molecular weight markers on sodium dodecyl sulfate polyacrylamide gel
CC electrophoresis techniques, for raising antibodies, and to test the
CC biological activities of the Ckb1 protein. (I)-HSA fusion proteins are
CC useful for screening for molecules that bind to the Ckb1 protein portion
CC of the fusion protein. The present sequence is used in the
CC exemplification of the present invention.

XX

SQ Sequence 656 AA;

Query Match 100.0%; Score 585; DB 7; Length 656;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRRFKDLGEENFKALVLIAPQVLCQCFEDHVKLVNEVTEFAKTCAVEDSAE 60
DB 72 DAHKEVAHRRFKDLGEENFKALVLIAPQVLCQCFEDHVKLVNEVTEFAKTCAVEDSAE 131
QY 61 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 132 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 191
QY 121 DVMTCTAFHDNEETFLKXYLVEIARRHPVEYAPPELLFPKRYKAAFTCCCAADKAAACLLP 180
DB 192 DVMTCTAFHDNEETFLKXYLVEIARRHPVEYAPPELLFPKRYKAAFTCCCAADKAAACLLP 251
QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFKAEFAEVSKLVTDLTK 240
DB 252 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFKAEFAEVSKLVTDLTK 311
QY 241 VHTCECHGDLLECADDDRADLAKYICENQDSISSKLKECCCKPILLESKSHCTAEVENDMPA 300
DB 312 VHTCECHGDLLECADDDRADLAKYICENQDSISSKLKECCCKPILLESKSHCTAEVENDMPA 371
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYBYARRHPDYSVVLRLRLAKTYETTLEK 360
DB 372 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYBYARRHPDYSVVLRLRLAKTYETTLEK 431
QY 361 CAADPHCEYAKVDFEFKPLVEEPONLIKONCELFQELGEYKFQNALLVRYTKVPQVST 420
DB 432 CAADPHCEYAKVDFEFKPLVEEPONLIKONCELFQELGEYKFQNALLVRYTKVPQVST 491
QY 421 PTLVSVSNLGVSKGCKCKHPEAKMPCAEYDLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 492 PTLVSVSNLGVSKGCKCKHPEAKMPCAEYDLSVLNQLCVLHEKTPVSDRVTKCCTES 551
QY 481 LVNRPFCFSALEVDVTVYPKGFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540
DB 552 LVNRPFCFSALEVDVTVYPKGFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 611
QY 541 KEQLKAVMDDFPAAFEVKCKCKADDDKTCFAEBGKGLVAASQAALGL 585
DB 612 KEQLKAVMDDFPAAFEVKCKCKADDDKTCFAEBGKGLVAASQAALGL 656

RESULT 42	
ADDC16791	
ADDC16791	standard; protein; 659 AA.
ADDC16791	
ADDC16791	
18-DEC-2003	(first entry)
Chimeric N-terminal T-20 (GGS)4GG-albumin fusion protein.	
human; anti-retroviral; T-20; T-1249; 5-Helix; cyanovirin-N; env; gp41;	
anti-HIV; vaccine; albumin fusion protein; HIV fusion inhibiting peptide;	
serum albumin; T-20-(GGS)4GG-albumin fusion; chimeric.	
Chimeric.	
Synthetic.	
Homo sapiens.	
Human immunodeficiency virus 1.	
Key	Location/Qualifiers
Peptide	1..24
/label= Signal_peptide	
Protein	25..659
/note= "Mature T-20-(GGS)4GG-albumin fusion protein"	
W0200306078-A1.	
14-AUG-2003.	
07-FEB-2003; 2003WO-1B000434.	
07-FEB-2002; 2002US-0355547P.	
(AVET) AVENTIS BEHRING GMBH.	
(DELZ) DELTA BIOTECHNOLOGY LTD.	
Hauser H, Weimer T, Sleep D;	
WPI; 2003-731478/69.	
N-PSDB; ADC16790.	
New albumin fusion protein comprising a human immunodeficiency virus	
(HIV) fusion inhibiting peptide and an albumin having an albumin	
activity, useful for treating a disease or disorder, e.g. HIV infection.	
Disclosure; Fig 6; 105pp; English.	
This invention relates to novel albumin fusion proteins comprising a	
human immunodeficiency virus (HIV) fusion inhibiting peptide, which	
exhibit anti-retroviral activity. Specifically, it refers to inhibitory	
peptides including T-20, T-1249, 5-Helix or cyanovirin-N that bind the	
HIV env protein, or derivatives thereof such as the HIV gp41 protein.	
Furthermore, the albumin activity has the ability to prolong the in vivo	
half-life of these HIV fusion inhibiting peptides. Accordingly, the	
present invention describes fusion proteins that neutralise HIV in a host	
by raising an immune response and also antibodies that inhibit viral	
infection of uninfected cells. In this way, a method exists to prevent,	
treat or ameliorate HIV infection and/ or a disease caused by HIV	
infection. As such, these composition have been described as having anti-	
HIV activity and can be used towards the production of a vaccine. This	
polypeptide sequence is the chimeric T-20-(GGS)4GG-albumin fusion protein	
of the invention.	
CC CC	
CC CC	
XX XX	
Sequence 659 AA;	
Query Match	100.0%; Score 585; DB 7; Length 659;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 585; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
1 DAHKSEVAHRFKD LG EENFKALVLI AFAQY LQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60	
75 DAHKSEVAHRFKD LG EENFKALVLI AFAQY LQCCPPEDHVKLVNEVTEFAKTCVADESAAE 134	

PI Hauser H, Weimer T, Sleep D;
XX WPI; 2003-731478/69.
DR N-PSDB; ADC16792.
XX
PT New albumin fusion protein comprising a human immunodeficiency virus
PT (HIV) fusion inhibiting peptide and an albumin having an albumin
PT activity, useful for treating a disease or disorder, e.g. HIV infection.
XX
PS Disclosure; Fig 8; 105pp; English.
XX
XX This invention relates to novel albumin fusion proteins comprising a
CC human immunodeficiency virus (HIV) fusion inhibiting peptide, which
CC exhibit anti-retroviral activity. Specifically, it refers to inhibitory
CC peptides including T-20, T-1249, 5-Helix or cyanovirin-N that bind the
CC HIV env protein, or derivatives thereof such as the HIV gp41 protein.
CC Furthermore, the albumin activity has the ability to prolong the in vivo
CC half-life of these HIV fusion inhibiting peptides. Accordingly, the
CC present invention describes fusion proteins that neutralise HIV in a host
CC by raising an immune response and also antibodies that inhibit viral
CC infection of uninfected cells. In this way, a method exists to prevent,
CC treat or ameliorate HIV infection and/or a disease caused by HIV
CC infection. As such, these composition have been described as having anti-
CC HIV activity and can be used towards the production of a vaccine. This
CC polypeptide sequence is the chimeric C-terminal albumin-(GGS)4GG-T-20
CC fusion protein of the invention.
XX
XX
SQ Sequence 659 AA;
Query Match 100.0%; Score 585; DB 7; Length 659;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRFKDGLGENFKALVLAFAQYLOQCPFDHVKLVNVEVTFKTCVADESAAE 60
DB 25 DAHKSEVAHRFKDGLGENFKALVLAFAQYLOQCPFDHVKLVNVEVTFKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRTYTGEMADCCAKQBPENECFLQHKDDNPNLPLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRTYTGEMADCCAKQBPENECFLQHKDDNPNLPLVRPEV 144
QY 121 DVNCTAFHNEETFLKKLYIEARRHPYFAPDELLFFAKRYKAAFTCCQAAADKAACLLP 180
DB 145 DVNCTAFHNEETFLKKLYIEARRHPYFAPDELLFFAKRYKAAFTCCQAAADKAACLLP 204
QY 181 KLDELDEGKASSAKORLKCASLQKGFERAFKAWAVARLSQFPKAEFAEYSKLVTLTK 240
DB 205 KLDELDEGKASSAKORLKCASLQKGFERAFKAWAVARLSQFPKAEFAEYSKLVTLTK 264
QY 241 VHTCCCHGDLLECADRDADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDENMPA 300
DB 265 VHTCCCHGDLLECADRDADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDENMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARRHPDYDYSVLLLLAKTYETTLK 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARRHPDYDYSVLLLLAKTYETTLK 384
QY 361 CAAADPHECYAKVDFDEKPLVPEEPQNLIKONCELPQELGEYKFQNALIVRYTKYPQVST 420
DB 385 CAAADPHECYAKVDFDEKPLVPEEPQNLIKONCELPQELGEYKFQNALIVRYTKYPQVST 444
QY 421 PTLVEVSRLNIGKVGSKCKRPEAKRMPCAEDYLSVNLQVLVHKTVPVSDEVTKCCCTES 480
DB 445 PTLVEVSRLNIGKVGSKCKRPEAKRMPCAEDYLSVNLQVLVHKTVPVSDEVTKCCCTES 504
QY 481 LVNRRPCFSALEVDETVPKPEFNAETFTPHADICTLSEKEROIKQTALVELVGHKPKAT 540
DB 505 LVNRRPCFSALEVDETVPKPEFNAETFTPHADICTLSEKEROIKQTALVELVGHKPKAT 564
QY 541 KEQJKAVMDDFAAFVEKCKKADKKECTCFABEGKKLVAAASQAALGL 585
DB 565 KEQJKAVMDDFAAFVEKCKKADKKECTCFABEGKKLVAAASQAALGL 609

RESULT 44
ADC16789
ID ADC16789 standard; protein; 662 AA.
XX
XX
AC ADC16789;
XX
DT 18-DEC-2003 (first entry)
XX
DE Chimeric C-terminal albumin-(GGS)4GG-T-1249 fusion protein sequence.
XX
KW human; anti-retroviral; T-20; T-1249; 5-Helix; cyanovirin-N; env; gp41;
KW anti-HIV; vaccine; albumin fusion protein; HIV fusion inhibiting peptide;
KW ds; serum albumin; gene; albumin-(GGS)4GG-T-1249 fusion; chimeric.
XX
OS Chimeric.
OS Synthetic.
OS Homo sapiens.
OS Human immunodeficiency virus 1.
XX
XX Key Location/Qualifiers
FH Peptide 1..24
FT /label= Signal_peptide
FT Protein 25..662
FT /note= "Mature albumin-(GGS)4GG-T-1249 fusion protein"
XX
PN WO2003066078-A1.
XX
XX
PD 14-AUG-2003.
XX
PF 07-FEB-2003; 2003WO-IB000434.
XX
XX 07-FEB-2002; 2002US-0355547P.
PR
XX (AVET) AVENTIS BEHRING GMBH.
PA (DELZ) DELTA BIOTECHNOLOGY LTD.
XX
XX Hauser H, Weimer T, Sleep D;
PI
XX WPI; 2003-731478/69.
DR
DR N-PSDB; ADC16788.
XX
XX New albumin fusion protein comprising a human immunodeficiency virus
PT (HIV) fusion inhibiting peptide and an albumin having an albumin
PT activity, useful for treating a disease or disorder, e.g. HIV infection.
XX
XX Disclosure; Fig 4; 105pp; English.
XX
XX This invention relates to novel albumin fusion proteins comprising a
CC human immunodeficiency virus (HIV) fusion inhibiting peptide, which
CC exhibit anti-retroviral activity. Specifically, it refers to inhibitory
CC peptides including T-20, T-1249, 5-Helix or cyanovirin-N that bind the
CC HIV env protein, or derivatives thereof such as the HIV gp41 protein.
CC Furthermore, the albumin activity has the ability to prolong the in vivo
CC half-life of these HIV fusion inhibiting peptides. Accordingly, the
CC present invention describes fusion proteins that neutralise HIV in a host
CC by raising an immune response and also antibodies that inhibit viral
CC infection of uninfected cells. In this way, a method exists to prevent,
CC treat or ameliorate HIV infection and/or a disease caused by HIV
CC infection. As such, these composition have been described as having anti-
CC HIV activity and can be used towards the production of a vaccine. This
CC polypeptide sequence is the chimeric C-terminal albumin-(GGS)4GG-T-1249
CC fusion protein of the invention.
XX
XX
SQ Sequence 662 AA;
Query Match 100.0%; Score 585; DB 7; Length 662;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRFKDGLGENFKALVLAFAQYLOQCPFDHVKLVNVEVTFKTCVADESAAE 60
DB 25 DAHKSEVAHRFKDGLGENFKALVLAFAQYLOQCPFDHVKLVNVEVTFKTCVADESAAE 84

```
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 144
QY 121 DVMCTAFHDNEETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
DB 145 DVMCTAFHDNEETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRFPAEFAEYSLVTLDTK 240
DB 205 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRFPAEFAEYSLVTLDTK 264
QY 241 VHTCCGDLLECCADRDADLAKYICENQDSISSKLKECEKPELLEKSHCIAEVENDEMPA 300
DB 265 VHTCCGDLLECCADRDADLAKYICENQDSISSKLKECEKPELLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEYARRHPDYSVLLLLSLAKTYETTLK 360
DB 325 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEYARRHPDYSVLLLLSLAKTYETTLK 384
QY 361 CAAADPHECVAKVDFDFKPLVBEPPQNLKQNCLEFEQGEYKFNQALLVRYTKKVPQVST 420
DB 385 CAAADPHECVAKVDFDFKPLVBEPPQNLKQNCLEFEQGEYKFNQALLVRYTKKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDETYVPKFNFAETFTFHADICTLSEKERQIKQTALVELVHKPKAT 540
DB 505 LVNRRPCFSALEVDETYVPKFNFAETFTFHADICTLSEKERQIKQTALVELVHKPKAT 564
QY 541 KEQLKAVMDDFAAAFVEKCKCKADDKETCFPAEKGKLVAAASQAALGL 585
DB 565 KEQLKAVMDDFAAAFVEKCKCKADDKETCFPAEKGKLVAAASQAALGL 609

RESULT 45
ACD16787
ID ACD16787 standard; protein; 662 AA.
XX
AC ACD16787;
XX
DT 18-DEC-2003 (first entry)
XX
DE Chimeric N-terminal T-1249 (GGS) 4GG-albumin fusion protein.
XX
KW human; anti-retroviral; T-20; T-1249; 5-Helix; cyanovirin-N; env; gp41;
KW anti-HIV; vaccine; albumin fusion protein; HIV fusion inhibiting peptide;
KW serum albumin; T-1249 (GGS) 4GG-albumin fusion; chimeric.
XX
OS Chimeric.
OS Synthetic.
OS Homo sapiens.
XX
Key Location/Qualifiers
FT Peptide 1..24
FT /label= Signal_peptide
FT Protein 25..662
FT /note= "Mature T-1249 (GGS) 4GG-albumin fusion protein"
XX
WO200306078-A1.
XX
14-AUG-2003.
XX
07-FEB-2003; 2003WO-IB000434.
XX
07-FEB-2002; 2002US-0355547P.
XX
(AVET ) AVENTIS BEHRING GMBH.
PA (DELZ ) DELTA BIOTECHNOLOGY LTD.
```

```
XX Hauser H, Weimer T, Sleep D;
PI
XX WPI. 2003-731478/69.
DR N-PSDB; ACD16786.
XX
PT New albumin fusion protein comprising a human immunodeficiency virus
PT (HIV) fusion inhibiting peptide and an albumin having an albumin
PT activity, useful for treating a disease or disorder, e.g. HIV infection.
XX
XX Disclosure; Fig 2; 105pp; English.
XX
CC This invention relates to novel albumin fusion proteins comprising a
CC human immunodeficiency virus (HIV) fusion inhibiting peptide, which
CC exhibit anti-retroviral activity. Specifically, it refers to inhibitory
CC peptides including T-20, T-1249, 5-Helix or cyanovirin-N that bind the
CC HIV env protein, or derivatives thereof such as the HIV gp41 protein.
CC Furthermore, the albumin activity has the ability to prolong the in vivo
CC half-life of these HIV fusion inhibiting peptides. Accordingly, the
CC present invention describes fusion proteins that neutralise HIV in a host
CC by raising an immune response and also antibodies that inhibit viral
CC infection of uninfected cells. In this way, a method exists to prevent,
CC treat or ameliorate HIV infection and/ or a disease caused by HIV
CC infection. As such, these compositions have been described as having anti-
CC HIV activity and can be used towards the production of a vaccine. This
CC polypeptide sequence is the chimeric T-1249 (GGS) 4GG-albumin fusion
CC protein of the invention.
XX
XX Sequence 662 AA;
XX
Query March 100.0%; Score 585; DB 7; Length 662;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAYIQCCPFEDHVKLVNVEVTEFAKTCVADSAE 60
DB 78 DAHKEVAHRFKDLGEENFKALVLIAPAYIQCCPFEDHVKLVNVEVTEFAKTCVADSAE 137
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
DB 138 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 197
QY 121 DVMCTAFHDNEETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
DB 198 DVMCTAFHDNEETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 257
QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRFPAEFAEYSLVTLDTK 240
DB 258 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRFPAEFAEYSLVTLDTK 317
QY 241 VHTCCGDLLECCADRDADLAKYICENQDSISSKLKECEKPELLEKSHCIAEVENDEMPA 300
DB 318 VHTCCGDLLECCADRDADLAKYICENQDSISSKLKECEKPELLEKSHCIAEVENDEMPA 377
QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEYARRHPDYSVLLLLSLAKTYETTLK 360
DB 378 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEYARRHPDYSVLLLLSLAKTYETTLK 437
QY 361 CAAADPHECVAKVDFDFKPLVBEPPQNLKQNCLEFEQGEYKFNQALLVRYTKKVPQVST 420
DB 438 CAAADPHECVAKVDFDFKPLVBEPPQNLKQNCLEFEQGEYKFNQALLVRYTKKVPQVST 497
QY 421 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 498 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 557
QY 481 LVNRRPCFSALEVDETYVPKFNFAETFTFHADICTLSEKERQIKQTALVELVHKPKAT 540
DB 558 LVNRRPCFSALEVDETYVPKFNFAETFTFHADICTLSEKERQIKQTALVELVHKPKAT 617
QY 541 KEQLKAVMDDFAAAFVEKCKCKADDKETCFPAEKGKLVAAASQAALGL 585
DB 618 KEQLKAVMDDFAAAFVEKCKCKADDKETCFPAEKGKLVAAASQAALGL 662
```

```

326 VHTCCHGDLLECADRADLAKYICENQDSISSKLEKCECEKPLEKSHGICAEVENDMPA 385
301 DLPSLAADFVESKDVCKVYAEAKDVLGMFLYVYARRHPDYSVLLRLAKTVETLEKC 360
386 DLPSLAADFVESKDVCKVYAEAKDVLGMFLYVYARRHPDYSVLLRLAKTVETLEKC 445
361 CAAADPHCYAKVDFEFKPLVEEPQNLIKONCBELFQOLGEYKFNQALLVRYTKKVPQVST 420
446 CAAADPHCYAKVDFEFKPLVEEPQNLIKONCBELFQOLGEYKFNQALLVRYTKKVPQVST 505
421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTES 480
506 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTES 565
481 LVNRRPCFSALEVDETVVPKFEFNAETFTFHADICTLSEKERQIKKQTALVELVHKHPKAT 540
566 LVNRRPCFSALEVDETVVPKFEFNAETFTFHADICTLSEKERQIKKQTALVELVHKHPKAT 625
541 KEQLKAVMDDFAAFVEKCKADDKETCFABEGKKLVAAASQAALGL 585
626 KEQLKAVMDDFAAFVEKCKADDKETCFABEGKKLVAAASQAALGL 670

RESULT 47
AAB36550
ID AAB36550 standard; protein; 670 AA.
XX
AC AAB36550;
XX
DT 07-MAR-2001 (first entry)
XX
DE Recombinant human serum albumin (HSA) protein sequence #2.
XX
KW Human serum albumin; HSA.
XX
OS Homo sapiens.
XX
PN CN1266100-A.
XX
PD 13-SEP-2000.
XX
PF 04-MAR-1999; 99CN-00102794.
XX
PR 04-MAR-1999; 99CN-00102794.
XX
PA (MAOJ-) MAOJI BIOLOGICAL ENG SCI & TECH CO LTD.
XX
PI Liu Z;
XX
DR WPI; 2000-673207/66.
XX
PT Novel methods for the chemical synthesis, expression and recombinant
protein production for human serum albumin reformed gene.
XX
PS Example 1; Fig 4; 85pp; Chinese.
XX
CC The present invention relates to two kinds of DNA sequences of coded
human serum albumin (HSA), i.e. design of structure-modified gene segment
of HSA and artificial total synthesis and a production process for large-
scale production of genetic recombinant HSA by using methanol, yeast and
engineering bacterium, and discovers that the structure-modified gene can
greatly increase the expression quantity of HSA. The production process
can make the structural gene of HSA obtain high-level expression under
the drive of promoter induced by methanol, and make the HSA expression
product secrete into the fermenting liquor culture medium, and provide
reliable test data for more large-scale pilot-amplification of gene
engineering HSA. The present sequence represents a recombinant HSA
protein from the present invention
XX
SQ Sequence 670 AA;

Query Match 100.0%; Score 585; DB 3; Length 670;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDGLGGENFKALVLIAPAOYLQOCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 86 DAHKSEVAHRFKDGLGGENFKALVLIAPAOYLQOCPFEDHVKLVNEVTEFAKTCVADESAAE 145
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVRPEV 120
DB 146 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVRPEV 205
QY 121 DVNCTAFHDNEETFLKKLYIEIARRHPYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 206 DVNCTAFHDNEETFLKKLYIEIARRHPYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 265
QY 181 KLDELDRGKASAKQRLKACSLQKGERAFKAWAVARLSQRPFKABEAVSKLVTDLTK 240
DB 266 KLDELDRGKASAKQRLKACSLQKGERAFKAWAVARLSQRPFKABEAVSKLVTDLTK 325
QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLEKCECEKPLEKSHGICAEVENDMPA 300

```

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRRFKDLGSENFKALVLIAPFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAB 60
 Db 86 DAHKEVAHRRFKDLGSENFKALVLIAPFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAB 145
 QY 61 NCKSLHTLFGDKLCTVATLETTYGENADCCAKQEPERNECFLOHKDDNPPLRLVRPEV 120
 Db 146 NCKSLHTLFGDKLCTVATLETTYGENADCCAKQEPERNECFLOHKDDNPPLRLVRPEV 205
 QY 121 DVNCTAFHDNBEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
 Db 206 DVNCTAFHDNBEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 265
 QY 181 KLDELDRSGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
 Db 266 KLDELDRSGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 325
 QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVNDENPA 300
 Db 326 VHTCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVNDENPA 385
 QY 301 DLPSLAADFVSKDVKCKNYAEAKDVFLGMFLYIARRHPDYSVLLRLAKTYETTLK 360
 Db 386 DLPSLAADFVSKDVKCKNYAEAKDVFLGMFLYIARRHPDYSVLLRLAKTYETTLK 445
 QY 361 CAADPHCYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420
 Db 446 CAADPHCYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 505
 QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 Db 506 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 565
 QY 481 LVNRRPCFSALEVDITYVPKEFNAETTFHADICTLSEKERQIKKOTALVELVGHKPKAT 540
 Db 566 LVNRRPCFSALEVDITYVPKEFNAETTFHADICTLSEKERQIKKOTALVELVGHKPKAT 625
 QY 541 KEQLKAVNDDFAAFVEKCKCKADKXETCFABEGKLVAAASQAALGL 595
 Db 626 KEQLKAVNDDFAAFVEKCKCKADKXETCFABEGKLVAAASQAALGL 670

RESULT 48
 ABR42610
 ID ABR42610 standard; protein; 672 AA.
 AC ABR42610;
 DT 23-OCT-2003 (revised)
 DT 26-AUG-2003 (first entry)
 XX Fusion protein of abrogen and human serum albumin.
 DE Human; abrogen; kringling; hATF-kringle; angiogenesis; inhibitor;
 XX urokinase plasminogen activator; tumour; metastasis; cytostatic;
 KW gene therapy; serum albumin.
 XX Homo sapiens.
 OS Chimeric.
 XX
 XX Key Location/Qualifiers
 FT 2..87
 FT Region /note= "human abrogen"
 FT 88..672
 FT Region /note= "human serum albumin"
 FT
 XX WO2003042354-A2.
 XX
 XX 22-MAY-2003.
 PD
 XX 04-SEP-2002; 2002WO-US027895.
 XX
 XX

PR 04-SEP-2001; 2001US-0316300P.
 XX (AVET) AVENTIS PHARM INC.
 XX Resbit M, Fong TC, Brockstedt D;
 XX WPI; 2003-449566/42.
 XX New abrogen polypeptide, useful for treating an angiogenesis related
 PT diseases e.g. tumor metastasis.
 XX Example 8; Page 27; 95pp; English.
 PS The present sequence is the protein sequence of a fusion protein
 CC comprising human serum albumin and (N-terminally) a novel human abrogen,
 CC designated hATF-kringle, comprising the human urokinase plasminogen
 CC activator kringling domain. Abrogens such as hATF-kringle are potent
 CC inhibitors of endothelial proliferation and angiogenesis. Abrogen
 CC polypeptides are capable of inhibiting or reducing cell proliferation
 CC induced by both basic fibroblast growth factor (bFGF) and vascular
 CC endothelial growth factor in a specific endothelial cell proliferation
 CC assay; angiotensin only inhibits bFGF induced proliferation in this
 CC assay. Vectors that expressed abrogen polypeptides in vivo were shown to
 CC reduce tumour metastasis in 2 lung cancer models. The invention provides
 CC abrogen polypeptides and polynucleotides, and methods of using these to
 CC treat an angiogenesis-related disease or disorder, e.g. tumour metastasis
 CC (claimed). The abrogens may be produced as fusion proteins including e.g.
 CC an N-terminal interleukin-2 signal peptide and a C-terminal stabilising
 CC molecule such as human serum albumin. (Updated on 23-OCT-2003 to
 CC standardise OS field)
 XX Sequence 672 AA;
 XX
 Query Match 100.0%; Score 585; DB 7; Length 672;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRRFKDLGSENFKALVLIAPFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAB 60
 Db 86 DAHKEVAHRRFKDLGSENFKALVLIAPFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAB 147
 QY 61 NCKSLHTLFGDKLCTVATLETTYGENADCCAKQEPERNECFLOHKDDNPPLRLVRPEV 120
 Db 146 NCKSLHTLFGDKLCTVATLETTYGENADCCAKQEPERNECFLOHKDDNPPLRLVRPEV 207
 QY 121 DVNCTAFHDNBEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
 Db 206 DVNCTAFHDNBEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 267
 QY 181 KLDELDRSGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
 Db 266 KLDELDRSGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 327
 QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVNDENPA 300
 Db 328 VHTCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVNDENPA 387
 QY 301 DLPSLAADFVSKDVKCKNYAEAKDVFLGMFLYIARRHPDYSVLLRLAKTYETTLK 360
 Db 388 DLPSLAADFVSKDVKCKNYAEAKDVFLGMFLYIARRHPDYSVLLRLAKTYETTLK 447
 QY 361 CAADPHCYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420
 Db 448 CAADPHCYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 507
 QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 Db 508 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 567
 QY 481 LVNRRPCFSALEVDITYVPKEFNAETTFHADICTLSEKERQIKKOTALVELVGHKPKAT 540
 Db 568 LVNRRPCFSALEVDITYVPKEFNAETTFHADICTLSEKERQIKKOTALVELVGHKPKAT 627

QY 541 KEQLKAVMDPFAAFVEKCKADKCTCFABEGKLVAAASQAALGL 585
DB 628 KEQLKAVMDPFAAFVEKCKADKCTCFABEGKLVAAASQAALGL 672

RESULT 49

ABR42609
ID ABR42609 standard; protein; 674 AA.
AC ABR42609;
XX 23-OCT-2003 (revised)
DT 26-AUG-2003 (first entry)
XX Fusion protein of abrogen and human serum albumin.
DE Human; abrogen; kringle; mATP-kringle; angiogenesis; inhibitor;
KW urokinase plasminogen activator; tumour; metastasis; cytostatic;
KW gene therapy; serum albumin.
XX Homo sapiens.
OS Chimeric.

Key	Location/Qualifiers
Region	2..586
Region	/note= "human serum albumin"
Region	587..588
Region	/note= "peptide linker"
Region	589..674
Region	/note= "human abrogen"

WO2003042354-A2.

22-MAY-2003.

04-SEP-2002; 2002WO-US027885.

04-SEP-2001; 2001US-0316300P.

(AVET) AVENTIS PHARM INC.

Nesbit M, Fong TC, Brockstedt D;

WPI; 2003-449566/42.

New abrogen polypeptide, useful for treating an angiogenesis related diseases e.g. tumor metastasis.

Example 8; Page 27; 95pp; English.

The present sequence is the protein sequence of a fusion protein comprising human serum albumin joined via a peptide linker to a novel human abrogen, designated hATP-kringle, comprising the human urokinase plasminogen activator kringle domain. The sequence includes an N-terminal alanine residue that results from cleavage of an interleukin-2 signal peptide. Abrogens such as hATP-kringle are potent inhibitors of endothelial proliferation and angiogenesis. Abrogen polypeptides are capable of inhibiting or reducing cell proliferation induced by both basic fibroblast growth factor (bFGF) and vascular endothelial growth factor in a specific endothelial cell proliferation assay; angiostatin only inhibits bFGF induced proliferation in this assay. Vectors that expressed abrogen polypeptides in vivo were shown to reduce tumour metastasis in 2 lung cancer models. The invention provides abrogen polypeptides and polynucleotides, and methods of using these to treat angiogenesis-related disease or disorder, e.g. tumour metastasis (claimed). The abrogens may be produced as fusion proteins including e.g. an N-terminal interleukin-2 signal peptide and a C-terminal stabilising molecule such as human serum albumin. (Updated on 23-OCT-2003 to standardise OS field)

Sequence 674 AA;

Query Match 100.0%; Score 585; DB 7; Length 674;

Best Local Similarity	100.0%;	Pred. No. 0;	Matches	585;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	DAKSEVAHRFKDLGEENFKALVLIATAFYLOQCFFEDHVKLVNVTFFAKTCVADESAE	60									
DB	2	DAKSEVAHRFKDLGEENFKALVLIATAFYLOQCFFEDHVKLVNVTFFAKTCVADESAE	61									
QY	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPFRNECFLOHKDDNPNLPRLVRPEV	120									
DB	62	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPFRNECFLOHKDDNPNLPRLVRPEV	121									
QY	121	DVMCTAFHDNEETFLKKYLYETARRHPYFAYAPPELLFAYAKRYKAAFTCCQAAADKKAACLLP	180									
DB	122	DVMCTAFHDNEETFLKKYLYETARRHPYFAYAPPELLFAYAKRYKAAFTCCQAAADKKAACLLP	181									
QY	181	KLDELDEGKASSAKQRLKASIQKFGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK	240									
DB	182	KLDELDEGKASSAKQRLKASIQKFGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK	241									
QY	241	VHTECHGDLLECADRADLAKYICENQDSISKLKECCCKPILLESKSHCIAEVENDMPA	300									
DB	242	VHTECHGDLLECADRADLAKYICENQDSISKLKECCCKPILLESKSHCIAEVENDMPA	301									
QY	301	DLPSLAADFVESKDVCKNYAEAKDVFGLMFLVEYARRHPDYSVLLLRLLAKTYETTLK	360									
DB	302	DLPSLAADFVESKDVCKNYAEAKDVFGLMFLVEYARRHPDYSVLLLRLLAKTYETTLK	361									
QY	361	CAAADPHECVAKVPDEFKPLVEEPQNLIKQNCBFLPOLGEYKFNALLVRYTKVPQVST	420									
DB	362	CAAADPHECVAKVPDEFKPLVEEPQNLIKQNCBFLPOLGEYKFNALLVRYTKVPQVST	421									
QY	421	PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTES	480									
DB	422	PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTES	481									
QY	481	LVNRRPCFSALAEVDVETVPKEFNAETTFHADICTLSEKERQIKKQTALVELVKHKPKAT	540									
DB	482	LVNRRPCFSALAEVDVETVPKEFNAETTFHADICTLSEKERQIKKQTALVELVKHKPKAT	541									
QY	541	KEQLKAVMDPFAAFVEKCKADKCTCFABEGKLVAAASQAALGL	585									
DB	542	KEQLKAVMDPFAAFVEKCKADKCTCFABEGKLVAAASQAALGL	586									

RESULT 50

ADD06591

ID ADD06591 standard; protein; 676 AA.

XX AC ADD06591;

XX DT 01-JAN-2004 (first entry)

Human Ckb1-HAS fusion protein construct protein SEQ ID NO:127.

human; chemokine beta1; Ckb1; anti-HIV; neuroprotective; antithyroid; antiarthritic; antirheumatic; immunosuppressive; nootropic; antiinflammatory; antiasthmatic; virucide; antiatherosclerotic; nephrotrophic; tuberculostatic; HIV; immune disorder; haematopoietic disorder; antimicrobial; infection; multiple sclerosis; Grave's disease; arthritis; autoimmune disorder; transplant rejection; neurodegenerative disorder; rheumatoid arthritis; inflammatory disease; asthma; allergic disorder; Alzheimer's disease; osteoarthritis; colitis; inflammatory bowel disease; glomerulonephritis; infectious disease; inflammatory kidney disease; hepatitis infection; herpes viral infection; tuberculosis; hepatitis infection; proliferative disorder; atherosclerosis; human serum albumin; HSA.

XX Synthetic.

OS Homo sapiens.

XX PN WO200297038-A2.

XX

05-DEC-2002. 272 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRPKAEFAVSKLVDTLTK 331
24-MAY-2002; 2002WO-US016525. 241 VHTCCCHGDLLECADRADLAKYICENQDSISGLKCECKEPLLEKSHGIAVENDMPA 300
25-MAY-2001; 2001US-0293212P. 332 VHTCCCHGDLLECADRADLAKYICENQDSISGLKCECKEPLLEKSHGIAVENDMPA 391
(HUMA-) HUMAN GENOME SCI INC. 301 DLPSLAADFVESKDVCKVNAEAKDVLGMFLYEHARRHPDYSVLLRLAKTYETTLK 360
Bell A, Ruben SM; 392 DLPSLAADFVESKDVCKVNAEAKDVLGMFLYEHARRHPDYSVLLRLAKTYETTLK 451
WPI, 2003-140456/13. 361 CAAADPHECYAKVDFEPKPLVEEPQNLIKONCELFQOLGEYFQNALVRYTKYQPVST 420
N-PSDB; ADD06600. 452 CAAADPHECYAKVDFEPKPLVEEPQNLIKONCELFQOLGEYFQNALVRYTKYQPVST 511
Novel human chemokine betal protein comprising deletion in amino acids 421 PTLVEVSRNLGKVGSKCKEPEAKRMPCEADYLSVLNQLCVLHEKTPVSDRVTKCTES 480
from amino and/or carboxy terminus, and is a fusion protein further 512 PTLVEVSRNLGKVGSKCKEPEAKRMPCEADYLSVLNQLCVLHEKTPVSDRVTKCTES 571
comprising human serum albumin, is useful for treating multiple 481 LVNRRPCFSALEVDYVYPKEFNAETTFHADICTLSEKEROIKKQTALVELVGHKPKAT 540
sclerosis, asthma. 572 LVNRRPCFSALEVDYVYPKEFNAETTFHADICTLSEKEROIKKQTALVELVGHKPKAT 631
Example 1; SEQ ID NO 127; 423pp; English. 541 KEQLKAVMDPFAAFVEKCKADDDKEICPAEKGKLVAAQAALGL 585
The present invention describes a human chemokine betal (Ckb1) protein 632 KEQLKAVMDPFAAFVEKCKADDDKEICPAEKGKLVAAQAALGL 676
(I) comprising a deletion in amino acid residues from the amino terminus
and/or carboxy terminus of the 93 residue amino acid sequence (SI, see
ADD06466). (I) has anti-HIV, neuroprotective, antithyroid, antiarthritic,
antirheumatic, immunosuppressive, neutropenic, antiinflammatory,
antiasthmatic, antiallergic, osteopathic, nephrotrophic, tuberculostatic,
virucide, antiatherosclerotic and antimicrobial activities (I) is useful
for preventing infection, preferably viral (human immunodeficiency virus
(HIV)) infection, in a cell, by contacting the cell with (I). (I) is also
useful for treating a disease, such as HIV infection or immune disorders,
haematopoietic disorders, autoimmune disorders, multiple sclerosis,
Grave's disease, arthritis, rheumatoid arthritis, transplant rejection,
neurodegenerative disorders, Alzheimer's disease, inflammatory disease,
asthma, allergic disorders, inflammatory bowel disease, osteoarthritis,
colitis, inflammatory kidney diseases, glomerulonephritis, infectious
disease, tuberculosis, hepatitis infections, herpes viral infection,
viral infection, proliferative disorders or atherosclerosis, in an
individual. (I) inhibits or abolishes the ability of HIV to bind to,
enter into/fuse with (infect), and/or replicate in CCR5 expressing cells.
(I) also acts a CCR5 agonist or antagonist, stimulate chemotaxis of
CCR5-expressing cells, inhibit CCR5 ligand binding to a CCR5 molecule, or
upregulate or downregulate CCR5 expression. (I) is useful as an
immunological probe for the differential identification of the tissues or
cell-types. (I)-human serum albumin (HSA) fusion proteins are useful for
diagnosing, treating and preventing various disorders in mammals,
preferably in humans. (I)-HSA fusion proteins are also useful as
molecular weight markers on sodium dodecyl sulfate polyacrylamide gel
electrophoresis techniques, for raising antibodies, and to test the
biological activities of the Ckb1 protein. (I)-HSA fusion proteins are
useful for screening for molecules that bind to the Ckb1 protein portion
of the fusion protein. The present sequence is used in the
exemplification of the present invention.

PD XX 05-DEC-2002. 272 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRPKAEFAVSKLVDTLTK 331
XX XX 24-MAY-2002; 2002WO-US016525. 241 VHTCCCHGDLLECADRADLAKYICENQDSISGLKCECKEPLLEKSHGIAVENDMPA 300
XX XX 25-MAY-2001; 2001US-0293212P. 332 VHTCCCHGDLLECADRADLAKYICENQDSISGLKCECKEPLLEKSHGIAVENDMPA 391
XX XX (HUMA-) HUMAN GENOME SCI INC. 301 DLPSLAADFVESKDVCKVNAEAKDVLGMFLYEHARRHPDYSVLLRLAKTYETTLK 360
XX XX Bell A, Ruben SM; 392 DLPSLAADFVESKDVCKVNAEAKDVLGMFLYEHARRHPDYSVLLRLAKTYETTLK 451
XX XX WPI, 2003-140456/13. 361 CAAADPHECYAKVDFEPKPLVEEPQNLIKONCELFQOLGEYFQNALVRYTKYQPVST 420
XX XX N-PSDB; ADD06600. 452 CAAADPHECYAKVDFEPKPLVEEPQNLIKONCELFQOLGEYFQNALVRYTKYQPVST 511
XX XX Novel human chemokine betal protein comprising deletion in amino acids 421 PTLVEVSRNLGKVGSKCKEPEAKRMPCEADYLSVLNQLCVLHEKTPVSDRVTKCTES 480
XX XX from amino and/or carboxy terminus, and is a fusion protein further 512 PTLVEVSRNLGKVGSKCKEPEAKRMPCEADYLSVLNQLCVLHEKTPVSDRVTKCTES 571
XX XX comprising human serum albumin, is useful for treating multiple 481 LVNRRPCFSALEVDYVYPKEFNAETTFHADICTLSEKEROIKKQTALVELVGHKPKAT 540
XX XX sclerosis, asthma. 572 LVNRRPCFSALEVDYVYPKEFNAETTFHADICTLSEKEROIKKQTALVELVGHKPKAT 631
XX XX Example 1; SEQ ID NO 127; 423pp; English. 541 KEQLKAVMDPFAAFVEKCKADDDKEICPAEKGKLVAAQAALGL 585
XX XX The present invention describes a human chemokine betal (Ckb1) protein 632 KEQLKAVMDPFAAFVEKCKADDDKEICPAEKGKLVAAQAALGL 676
XX XX (I) comprising a deletion in amino acid residues from the amino terminus
XX XX and/or carboxy terminus of the 93 residue amino acid sequence (SI, see
XX XX ADD06466). (I) has anti-HIV, neuroprotective, antithyroid, antiarthritic,
XX XX antirheumatic, immunosuppressive, neutropenic, antiinflammatory,
XX XX antiasthmatic, antiallergic, osteopathic, nephrotrophic, tuberculostatic,
XX XX virucide, antiatherosclerotic and antimicrobial activities (I) is useful
XX XX for preventing infection, preferably viral (human immunodeficiency virus
XX XX (HIV)) infection, in a cell, by contacting the cell with (I). (I) is also
XX XX useful for treating a disease, such as HIV infection or immune disorders,
XX XX haematopoietic disorders, autoimmune disorders, multiple sclerosis,
XX XX Grave's disease, arthritis, rheumatoid arthritis, transplant rejection,
XX XX neurodegenerative disorders, Alzheimer's disease, inflammatory disease,
XX XX asthma, allergic disorders, inflammatory bowel disease, osteoarthritis,
XX XX colitis, inflammatory kidney diseases, glomerulonephritis, infectious
XX XX disease, tuberculosis, hepatitis infections, herpes viral infection,
XX XX viral infection, proliferative disorders or atherosclerosis, in an
XX XX individual. (I) inhibits or abolishes the ability of HIV to bind to,
XX XX enter into/fuse with (infect), and/or replicate in CCR5 expressing cells.
XX XX (I) also acts a CCR5 agonist or antagonist, stimulate chemotaxis of
XX XX CCR5-expressing cells, inhibit CCR5 ligand binding to a CCR5 molecule, or
XX XX upregulate or downregulate CCR5 expression. (I) is useful as an
XX XX immunological probe for the differential identification of the tissues or
XX XX cell-types. (I)-human serum albumin (HSA) fusion proteins are useful for
XX XX diagnosing, treating and preventing various disorders in mammals,
XX XX preferably in humans. (I)-HSA fusion proteins are also useful as
XX XX molecular weight markers on sodium dodecyl sulfate polyacrylamide gel
XX XX electrophoresis techniques, for raising antibodies, and to test the
XX XX biological activities of the Ckb1 protein. (I)-HSA fusion proteins are
XX XX useful for screening for molecules that bind to the Ckb1 protein portion
XX XX of the fusion protein. The present sequence is used in the
XX XX exemplification of the present invention.

Query Match 100.0%; Score 585; DB 7; Length 676;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKEVAHAFKDLGEENFKALVLAFAOVLQCPEDHVKLVNEVTEFAKTCVADESAAE 60
DB 92 DAHKEVAHAFKDLGEENFKALVLAFAOVLQCPEDHVKLVNEVTEFAKTCVADESAAE 151
QY 61 NCDKSLHTLFGDKLCTVATLRETYGENADCCAKQEPERNECFLOHKDNDPNLPRLVREPV 120
DB 152 NCDKSLHTLFGDKLCTVATLRETYGENADCCAKQEPERNECFLOHKDNDPNLPRLVREPV 211
QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFVAPPELLFFAKYKAAFTCCQADKAACLIP 180
DB 212 DVMCTAFHDNEETFLKYLVEIARRHPYFVAPPELLFFAKYKAAFTCCQADKAACLIP 271
QY 181 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRPKAEFAVSKLVDTLTK 240

CC colitis, inflammatory kidney diseases, glomerulonephritis, infectious
CC disease, tuberculosis, hepatitis infections, herpes viral infection,
CC viral infection, proliferative disorders or atherosclerosis, in an
CC individual (I) inhibits or abolishes the ability of HIV to bind to,
CC enter into/fuse with (infect), and/or replicate in CCR5 expressing cells.
CC (I) also acts a CCR5 agonists or antagonists, stimulate chemotaxis of
CC CCR5-expressing cells, inhibit CCR5 ligand binding to a CCR5 molecule, or
CC upregulate or downregulate CCR5 expression. (I) is useful as an
CC immunological probe for the differential identification of the tissues or
CC cell-types. (I)-human serum albumin (HSA) fusion proteins are useful for
CC diagnosing, treating and preventing various disorders in mammals,
CC preferably in humans. (I)-HSA fusion proteins are also useful as
CC molecular weight markers on sodium dodecyl sulfate polyacrylamide gel
CC electrophoresis techniques, for raising antibodies, and to test the
CC biological activities of the Cxbl protein. (I)-HSA fusion proteins are
CC useful for screening for molecules that bind to the Cxbl protein portion
CC of the fusion protein. The present sequence is used in the
CC exemplification of the present invention.

XX	Sequence 677 AA;
DE	Query Match
XX	Best Local Similarity 100.0%; Score 585; DB 7; Length 677;
XX	Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 DAHKSVAHRFKDGLGEENFKALVIAFAQYLOQCFFEDHVKLVNEVTEFAKTCVADESA 60
Db	93 DAHKSVAHRFKDGLGEENFKALVIAFAQYLOQCFFEDHVKLVNEVTEFAKTCVADESA 152
QY	61 NCDKSLHTLFGKCLCTVATLRTYEMADCCAKQEPERNECFLOKDDNPNPLRLVPEV 120
Db	153 NCDKSLHTLFGKCLCTVATLRTYEMADCCAKQEPERNECFLOKDDNPNPLRLVPEV 212
QY	121 DVMTAFHNEETFLKLYIETARRHPFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
Db	213 DVMTAFHNEETFLKLYIETARRHPFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 272
QY	181 KLDELDEGKASSAKQRLKCAQLQKGFGERAFKAWAVARLSQRPFAFVSKVLVTLTK 240
Db	273 KLDELDEGKASSAKQRLKCAQLQKGFGERAFKAWAVARLSQRPFAFVSKVLVTLTK 332
QY	241 VHTCECHGDLLECADRADLAKYICENQDSISSKLKCECEKPLEKSHCIAEVENDEMPA 300
Db	333 VHTCECHGDLLECADRADLAKYICENQDSISSKLKCECEKPLEKSHCIAEVENDEMPA 392
QY	301 DLPSLAADFVESKDVCNVAEAKDVLGMLFLYAYARRHPDYSVLLRLAKTYETTLK 360
Db	393 DLPSLAADFVESKDVCNVAEAKDVLGMLFLYAYARRHPDYSVLLRLAKTYETTLK 452
QY	361 CAAADPEHCYAKVDFBEKPLVEEPQNLKONCELEOLGEYKFQNALVRYTKVPQVST 420
Db	453 CAAADPEHCYAKVDFBEKPLVEEPQNLKONCELEOLGEYKFQNALVRYTKVPQVST 512
QY	421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db	513 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 572
QY	481 LVNRPFCFSALEVDVETVPKFEAFETTHADICTLSEKEROIKQTALVELVHKPKAT 540
Db	573 LVNRPFCFSALEVDVETVPKFEAFETTHADICTLSEKEROIKQTALVELVHKPKAT 632
QY	541 KEQLKAVNDDFAAFVEKCKADDKETCFABEGKLVAAQQAALGL 585
Db	633 KEQLKAVNDDFAAFVEKCKADDKETCFABEGKLVAAQQAALGL 677
RESULT 53	
ADD06587	
ID	ADD06587 standard; protein; 680 AA.
XX	
AC	ADD06587;
XX	
DT	01-JAN-2004 (first entry)

Human Cxbl-HSA fusion protein construct protein SEQ ID NO:123.
human; chemokine betal; Cxbl; anti-HIV; neuroprotective; antithyroid;
antiarthritic; antirheumatic; immunosuppressive; nootropic;
antiinflammatory; antiasthmatic; antiallergic; osteopathic;
nephrotrophic; tuberculostatic; virucide; antiatherosclerotic;
antimicrobial; infection; HIV; immune disorder; haematopoietic disorder;
autoimmune disorder; multiple sclerosis; Grave's disease; arthritis;
rheumatoid arthritis; transplant rejection; neurodegenerative disorder;
Alzheimer's disease; inflammatory disease; osteoarthritis; colitis;
inflammatory bowel disease; glomerulonephritis; infectious disease;
tuberculosis; hepatitis infection; herpes viral infection;
human serum albumin, HSA.
Synthetic.
Homo sapiens.
WO200297038-A2.
05-DEC-2002.
24-MAY-2002; 2002WO-US016525.
25-MAY-2001; 2001US-02933212P.
(HUMA-) HUMAN GENOME SCI INC.
Bell A, Ruben SM;
WPI; 2003-140456/13.
N-P8DB; ADD06598.
Novel human chemokine betal protein comprising deletion in amino acids
from amino and/or carboxy terminus, and is a fusion protein further
comprising human serum albumin, is useful for treating multiple
sclerosis, asthma.
Example 1; SEQ ID NO 123; 423pp; English.

The present invention describes a human chemokine betal (Cxbl) protein
(I) comprising a deletion in amino acid residues from the amino terminus
and/or carboxy terminus of the 93 residue amino acid sequence (S1, see
ADD06468). (I) has anti-HIV, neuroprotective, antithyroid, antiarthritic,
antiinflammatory, immunosuppressive, nootropic, antiinflammatory,
antiasthmatic, antiallergic, osteopathic, nephrotrophic, tuberculostatic,
virucide, antiatherosclerotic and antimicrobial activities. (I) is useful
for preventing infection, preferably viral (human immunodeficiency virus
(HIV)) infection, in a cell, by contacting the cell with (I). (I) is also
useful for treating a disease, such as HIV infection or immune disorders,
haematopoietic disorders, autoimmune disorders, multiple sclerosis,
Grave's disease, arthritis, rheumatoid arthritis, transplant rejection,
neurodegenerative disorders, Alzheimer's disease, inflammatory disease,
asthma, allergic disorders, inflammatory bowel disease, osteoarthritis,
colitis, inflammatory kidney diseases, glomerulonephritis, infectious
disease, tuberculosis, hepatitis infections, herpes viral infection,
viral infection, proliferative disorders or atherosclerosis, in an
individual (I) inhibits or abolishes the ability of HIV to bind to,
enter into/fuse with (infect), and/or replicate in CCR5 expressing cells.
(I) also acts a CCR5 agonists or antagonists, stimulate chemotaxis of
CCR5-expressing cells, inhibit CCR5 ligand binding to a CCR5 molecule, or
upregulate or downregulate CCR5 expression. (I) is useful as an
immunological probe for the differential identification of the tissues or
cell-types. (I)-human serum albumin (HSA) fusion proteins are useful for
diagnosing, treating and preventing various disorders in mammals,
preferably in humans. (I)-HSA fusion proteins are also useful as
molecular weight markers on sodium dodecyl sulfate polyacrylamide gel
electrophoresis techniques, for raising antibodies, and to test the
biological activities of the Cxbl protein. (I)-HSA fusion proteins are
useful for screening for molecules that bind to the Cxbl protein portion
of the fusion protein. The present sequence is used in the

CC exemplification of the present invention.
XX
SQ Sequence 680 AA;

Query Match 100.0%; Score 585; DB 7; Length 680;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLQOCFFEDHVKLVNEVTEFAKTCVADESAE 60
Db DAHSEVAHRFKDLGSENFKALVLIAPAYLQOCFFEDHVKLVNEVTEFAKTCVADESAE 155

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 215

QY 121 DVNCTAFHDNEETFLKKLYEYARRHPYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db DVNCTAFHDNEETFLKKLYEYARRHPYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 275

QY 181 KLDELDEGKASSAKQRLKASLQKFGRAFKANAVARLSQRFPAEFAVSKLVTDLTK 240
Db KLDELDEGKASSAKQRLKASLQKFGRAFKANAVARLSQRFPAEFAVSKLVTDLTK 335

QY 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLKECCERPLLEKSHCIAEVENDEMPA 300
Db VHTECCHGDLLECADRADLAKYICENQDSISSKLKECCERPLLEKSHCIAEVENDEMPA 395

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYAYARRHPDYSVLLRLAKTVETTLK 360
Db DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYAYARRHPDYSVLLRLAKTVETTLK 455

QY 361 CAADPHCEYAKVDFEFPKLVSEPNQNLKONCELPQGEYKFNALLVRYTKKVPQVST 420
Db CAADPHCEYAKVDFEFPKLVSEPNQNLKONCELPQGEYKFNALLVRYTKKVPQVST 515

QY 421 PTLVEYSRNLGVKSGKCKKHPKAMPKAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db PTLVEYSRNLGVKSGKCKKHPKAMPKAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 575

QY 481 LVNRRFCFSALEVDYVPEKEFNASTFTFHADICTLSEKEROIKKOTALVELVKHFKPAT 540
Db LVNRRFCFSALEVDYVPEKEFNASTFTFHADICTLSEKEROIKKOTALVELVKHFKPAT 635

QY 541 KEOLKAVMDDFAAFVEKCKKADKDKETCFAEAGKGLVAASQAALGL 585
Db KEOLKAVMDDFAAFVEKCKKADKDKETCFAEAGKGLVAASQAALGL 680

RESULT 54
ABR42612
ID ABR42612 standard; protein; 687 AA.
XX
AC ABR42612;
XX AC
DT 23-OCT-2003 (revised)
DT 26-AUG-2003 (first entry)
XX
DE Fusion protein of abrogen and human serum albumin.
XX Human; abrogen; kringle; hATF-kringle; angiogenesis; inhibitor;
KW urokinase plasminogen activator; tumour; metastasis; cytostatic;
KW gene therapy; serum albumin.
XX
OS Homo sapiens.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Region 2..87
FT /note= "human abrogen"
FT Region 88..102
FT /note= "peptide linker"
FT Region 103..687

/note= "human serum albumin"
WO2003042354-A2.
22-MAY-2003.
04-SEP-2002; 2002WO-US027885.
04-SEP-2001; 2001US-0316300P.
(AVET) AVENTIS PHARM INC.
Nesbit M, Fong TC, Brockstedt D;
WPI; 2003-449566/42.
New abrogen polypeptide, useful for treating an angiogenesis related diseases e.g. tumor metastasis.
Example 8; Page 28; 95pp; English.
The present sequence is the protein sequence of a fusion protein comprising human serum albumin and (N-terminally) a novel human abrogen, designated hATF-kringle, comprising the human urokinase plasminogen activator kringle domain. Abrogens such as hATF-kringle are potent inhibitors of endothelial proliferation and angiogenesis. Abrogen polypeptides are capable of inhibiting or reducing cell proliferation induced by both basic fibroblast growth factor (bFGF) and vascular endothelial growth factor in a specific endothelial cell proliferation assay; angiostatin only inhibits bFGF induced proliferation in this assay. Vectors that expressed abrogen polypeptides in vivo were shown to reduce tumour metastasis in 2 lung cancer models. The invention provides abrogen polypeptides and polynucleotides, and methods of using these to treat an angiogenesis-related disease or disorder, e.g. tumour metastasis (claimed). The abrogens may be produced as fusion proteins including e.g. an N-terminal interleukin-2 signal peptide and a C-terminal stabilising molecule such as human serum albumin. (Updated on 23-OCT-2003 to standardise OS field)
SQ Sequence 687 AA;

Query Match 100.0%; Score 585; DB 7; Length 687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLQOCFFEDHVKLVNEVTEFAKTCVADESAE 60
Db DAHSEVAHRFKDLGSENFKALVLIAPAYLQOCFFEDHVKLVNEVTEFAKTCVADESAE 162

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 222

QY 121 DVNCTAFHDNEETFLKKLYEYARRHPYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db DVNCTAFHDNEETFLKKLYEYARRHPYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 282

QY 181 KLDELDEGKASSAKQRLKASLQKFGRAFKANAVARLSQRFPAEFAVSKLVTDLTK 240
Db KLDELDEGKASSAKQRLKASLQKFGRAFKANAVARLSQRFPAEFAVSKLVTDLTK 342

QY 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLKECCERPLLEKSHCIAEVENDEMPA 300
Db VHTECCHGDLLECADRADLAKYICENQDSISSKLKECCERPLLEKSHCIAEVENDEMPA 402

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYAYARRHPDYSVLLRLAKTVETTLK 360
Db DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYAYARRHPDYSVLLRLAKTVETTLK 462

QY 361 CAADPHCEYAKVDFEFPKLVSEPNQNLKONCELPQGEYKFNALLVRYTKKVPQVST 420
Db CAADPHCEYAKVDFEFPKLVSEPNQNLKONCELPQGEYKFNALLVRYTKKVPQVST 522

421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
523 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 582
481 LVNRPCFSALEVDYTVVPKFEAEFTFHADICTLSEKEROIKQTALVELVGHKPKAT 540
583 LVNRPCFSALEVDYTVVPKFEAEFTFHADICTLSEKEROIKQTALVELVGHKPKAT 642
541 KEQLKAVMDDFAAAFVEKCKCKADDKTCFAEKGKLVAAASQAALGL 585
643 KEQLKAVMDDFAAAFVEKCKCKADDKTCFAEKGKLVAAASQAALGL 687

RESULT 55

ABR42613 standard; protein; 688 AA.

AC ABR42613;
XX
XX
DT 23-OCT-2003 (revised)
DT 26-AUG-2003 (first entry)
XX
XX Fusion protein of abrogen and human serum albumin.

XX Human; abrogen; kringle; hATF-kringle; angiogenesis; inhibitor;
KW urokinase plasminogen activator; tumour; metastasis; cytostatic;
KW gene therapy; serum albumin.

OS Homo sapiens.
OS Chimeric.

XX Location/Qualifiers
FH Key 2..585
FT Region /note= "human serum albumin"
FT Region 585..602
FT Region /note= "peptide linker"
FT Region 603..688
FT Region /note= "human abrogen"

PN WO2003042354-A2.

XX 22-MAY-2003.

XX 04-SEP-2002; 2002WO-US027885.

XX 04-SEP-2001; 2001US-0316300P.

XX (AVET) AVENTIS PHARM INC.

XX Nesbit M, Fong TC, Bröckstedt D;

XX WPI; 2003-449566/42.

XX New abrogen polypeptide, useful for treating an angiogenesis related
XX diseases e.g. tumor metastasis.

XX Example 8; Page 28-29; 95pp; English.

XX The present sequence is the protein sequence of a fusion protein
XX comprising human serum albumin and (N-terminally) a novel human abrogen,
XX designated hATF-kringle, comprising the human urokinase plasminogen
XX activator kringle domain. Abrogens such as hATF-kringle are potent
XX inhibitors of endothelial proliferation and angiogenesis. Abrogen
XX polypeptides are capable of inhibiting or reducing cell proliferation
XX induced by both basic fibroblast growth factor (bFGF) and vascular
XX endothelial growth factor in a specific endothelial cell proliferation
XX assay; angiotatin only inhibits bFGF induced proliferation in this
XX assay. Vectors that expressed abrogen polypeptides in vivo were shown to
XX reduce tumour metastasis in 2 lung cancer models. The invention provides
XX abrogen polypeptides and polynucleotides, and methods of using these to
XX treat an angiogenesis-related disease or disorder, e.g. tumour metastasis
XX (claimed). The abrogens may be produced as fusion proteins including e.g.
XX an N-terminal interleukin-2 signal peptide and a C-terminal stabilising

CC molecule such as human serum albumin. (Updated on 23-OCT-2003 to
CC standardise OS field)

XX Sequence 688 AA;

Query Match 100.0%; Score 585; DB 7; Length 688;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGKGFENFKALVLIFAQYLOQCFFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHRFKDLGKGFENFKALVLIFAQYLOQCFFEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKQPERNECFLOHKKDDNPNLPLVRREV 120
DB 61 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKQPERNECFLOHKKDDNPNLPLVRREV 120

QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQADKAACLLP 180
DB 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQADKAACLLP 180

QY 181 KLDELDEGKASSAKORLKCASLOKGERAFKAWAVARLSORFPKAEFAEVSCLVDTLTK 240
DB 181 KLDELDEGKASSAKORLKCASLOKGERAFKAWAVARLSORFPKAEFAEVSCLVDTLTK 240

QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPPLLEKSHCIAEVENDEMPA 300

QY 301 DLPSLAADFEVSKDVCNVAEAKDVFGLMFLYEYASRHPDYSVLLLLAKTYETTLK 360
DB 301 DLPSLAADFEVSKDVCNVAEAKDVFGLMFLYEYASRHPDYSVLLLLAKTYETTLK 360

QY 361 CAAADPHCYAKVDFEPKPLVEEPQNLKQNCCLPEQLGEYKFQNALIVRYTKVPQVST 420
DB 361 CAAADPHCYAKVDFEPKPLVEEPQNLKQNCCLPEQLGEYKFQNALIVRYTKVPQVST 420

QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

QY 481 LVNRPCFSALEVDYTVVPKFEAEFTFHADICTLSEKEROIKQTALVELVGHKPKAT 540
DB 481 LVNRPCFSALEVDYTVVPKFEAEFTFHADICTLSEKEROIKQTALVELVGHKPKAT 540

QY 541 KEQLKAVMDDFAAAFVEKCKCKADDKTCFAEKGKLVAAASQAALGL 585

DB 541 KEQLKAVMDDFAAAFVEKCKCKADDKTCFAEKGKLVAAASQAALGL 585

RESULT 56

ABR42608

ID ABR42608 standard; protein; 689 AA.

XX AC ABR42608;

XX DT 23-OCT-2003 (revised)

XX DT 26-AUG-2003 (first entry)

XX Fusion protein of abrogen and human serum albumin.

XX Human; abrogen; kringle; hATF-kringle; angiogenesis; inhibitor;
XX urokinase plasminogen activator; tumour; metastasis; cytostatic;
XX gene therapy; serum albumin.

XX Homo sapiens.

XX Chimeric.

XX Key Location/Qualifiers

XX Region 2..586

XX FT /note= "human serum albumin"

XX FT 587..603

XX FT /note= "peptide linker"

Region 504. .689
/note="human abrogen"
WO2003042354-A2.
22-MAY-2003.
04-SEP-2002; 2002WO-US027885.
04-SEP-2001; 2001US-0316300P.
(AVET) AVENTIS PHARM INC.
Nesbit M, Fong TC, Brockstedt D;
WPI; 2003-449566/42.
New abrogen polypeptide, useful for treating an angiogenesis related
diseases e.g. tumor metastasis.
Example 8; Page 27; 95pp; English.
The present sequence is the protein sequence of a fusion protein
comprising human serum albumin joined via a peptide linker to a novel
human abrogen, designated hATF-kringle, comprising the human urokinase
plasminogen activator kringle domain. The sequence includes an N-terminal
alanine residue that results from cleavage of an interleukin-2 signal
peptide. Abrogens such as hATF-kringle are potent inhibitors of
endothelial proliferation and angiogenesis. Abrogen polypeptides are
capable of inhibiting or reducing cell proliferation induced by both
basic fibroblast growth factor (bFGF) and vascular endothelial growth
factor in a specific endothelial cell proliferation assay; angiostatin
only inhibits bFGF induced proliferation in this assay. Vectors that
expressed abrogen polypeptides in vivo were shown to reduce tumour
metastasis in 2 lung cancer models. The invention provides abrogen
polypeptides and polynucleotides, and methods of using these to treat an
angiogenesis-related disease or disorder, e.g. tumour metastasis
(claimed). The abrogens may be produced as fusion proteins including e.g.
an N-terminal interleukin-2 signal peptide and a C-terminal stabilising
molecule such as human serum albumin. (Updated on 23-OCT-2003 to
standardise OS field)

Sequence 689 AA;
Query Match 100.0%; Score 585; DB 7; Length 689;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 DAHSEVAHFRKDLGEENFKALVLIAPQVLOQCPFDHVKLVNEVTEFAKTCVADESAE 60
2 DAHSEVAHFRKDLGEENFKALVLIAPQVLOQCPFDHVKLVNEVTEFAKTCVADESAE 61
61 NCDXSLHTLFGDKLCTVATLRETYGENMADCCAKQOEPRNECFLOHKDDNPNLPLRVREV 120
62 NCDXSLHTLFGDKLCTVATLRETYGENMADCCAKQOEPRNECFLOHKDDNPNLPLRVREV 121
121 DVMCTAFHDNEETLKKYLVEIARRHFFYAPELLFFAKRYKAAFTCCQQAADKKAACLLP 180
122 DVMCTAFHDNEETLKKYLVEIARRHFFYAPELLFFAKRYKAAFTCCQQAADKKAACLLP 181
181 KLDELREGKASSAKQRLKCAQLQKFGERAFKAWAVARLSQRPPKAFEAESKLVTDJTK 240
182 KLDELREGKASSAKQRLKCAQLQKFGERAFKAWAVARLSQRPPKAFEAESKLVTDJTK 241
241 VHTCCGGDLLEACDDRADLAKYICENQDISSKLKECKECPKLLKSHCIAEVENDEMPA 300
242 VHTCCGGDLLEACDDRADLAKYICENQDISSKLKECKECPKLLKSHCIAEVENDEMPA 301
301 DLPLSAADFVESKDYCKNYAEAKQVFLGMFLYFYARRHPDYSVVLRLRAKTYETTLK 360
302 DLPLSAADFVESKDYCKNYAEAKQVFLGMFLYFYARRHPDYSVVLRLRAKTYETTLK 361
361 CAADNDETCVAKNFEFVPIVFEESONI KQNCETFEQIGEVKFNQAI VEVKMKVDPVST 420

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHREFKDLGEENFKALVLIAPQYLOQCPEDEHVKLVNEVTEFAKTCVADSEAE 60
 DB 25 DAHSEVAHREFKDLGEENFKALVLIAPQYLOQCPEDEHVKLVNEVTEFAKTCVADSEAE 84
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRLVPEV 120
 DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRLVPEV 144
 QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFAPPELLFPAKRYKAAATECCQADYAAACLLP 180
 DB 145 DVMCTAFHDNEETFLKYLVEIARRHPYFAPPELLFPAKRYKAAATECCQADYAAACLLP 204
 QY 191 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFAEFAEVS KLVTDLTK 240
 DB 205 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFAEFAEVS KLVTDLTK 264
 QY 241 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKLKECCKEKLLSKSHCIAEVENDEMPA 300
 DB 265 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKLKECCKEKLLSKSHCIAEVENDEMPA 324
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRKAKTYETTLK 360
 DB 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRKAKTYETTLK 384
 QY 361 CAAADPHCEYAKVDFEPKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPOVST 420
 DB 395 CAAADPHCEYAKVDFEPKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPOVST 444
 QY 421 PTLVEVRNLGKVGSKCKKPEAKRMPCAEDYLSVVLNOLCVLHSEKTPVSDRVTKCCTES 480
 DB 445 PTLVEVRNLGKVGSKCKKPEAKRMPCAEDYLSVVLNOLCVLHSEKTPVSDRVTKCCTES 504
 QY 481 LVNRPCFSALEVDVETVPKPEFAETFTPHADICTLSEKERQIKKOTALVELVXHKPKAT 540
 DB 505 LVNRPCFSALEVDVETVPKPEFAETFTPHADICTLSEKERQIKKOTALVELVXHKPKAT 564
 QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKGLVAASQAALGL 585
 DB 565 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKGLVAASQAALGL 609

RESULT 58
 AAR39477
 ID AAR39477 standard; protein; 787 AA.
 AC AAR39477;
 XX
 XX 25-MAR-2003 (revised)
 DT 28-JAN-1994 (first entry)
 XX
 DE G-CSF- (Gly)4-HSA chimera encoded by pYGI301.
 XX
 XX Granulocyte colony stimulating factor; G-CSF; human serum albumin; HSA;
 KW chimera; fusion protein; leukopenia; leukaemia; transplantation;
 XX immune system; bone marrow; cancer.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 1. 24
 FT Peptide /note= "prepro region of HSA"
 FT Protein 25..198 G-CSF
 FT Region 199..202 /label= linker
 FT Protein 203..787 /label= HSA
 FT
 FT
 XX W09315211-A1.
 PN
 XX
 PD 05-AUG-1993.

28-JAN-1993; 93WO-FR000086.
 31-JAN-1992; 92FR-00001065.
 (RHON) RHONE POULENC RORER SA.
 Yeh P;
 WPI: 1993-258686/32.
 N-PSDB; AAQ45988.
 New granulocyte colony stimulating factor fusion proteins - contg.
 stabilising protein, for treating leukopenia, leukaemia, etc.
 Disclosure; Fig 5; 36pp; French.
 G-CSF fusion proteins are capable of maintaining G-CSF activity for long periods in vivo. HSA-G-CSF has lower activity than native G-CSF in vitro but comparable activity in vivo. The prod. may be used to treat diseases requiring an increase in granulocyte count and/or activity, esp. leukaemia and certain forms of leukaemia, or to stimulate the immune system during transplantation (e.g. of bone marrow) or after cancer chemotherapy. (Updated on 25-MAR-2003 to correct PN field.)
 Sequence 787 AA;
 Query Match 100.0%; Score 585; DB 2; Length 787;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHREFKDLGEENFKALVLIAPQYLOQCPEDEHVKLVNEVTEFAKTCVADSEAE 60
 DB 203 DAHSEVAHREFKDLGEENFKALVLIAPQYLOQCPEDEHVKLVNEVTEFAKTCVADSEAE 262
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRLVPEV 120
 DB 263 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPRLVPEV 322
 QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFAPPELLFPAKRYKAAATECCQADYAAACLLP 180
 DB 323 DVMCTAFHDNEETFLKYLVEIARRHPYFAPPELLFPAKRYKAAATECCQADYAAACLLP 382
 QY 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFAEFAEVS KLVTDLTK 240
 DB 383 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFAEFAEVS KLVTDLTK 442
 QY 241 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKLKECCKEKLLSKSHCIAEVENDEMPA 300
 DB 443 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKLKECCKEKLLSKSHCIAEVENDEMPA 502
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRKAKTYETTLK 360
 DB 503 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRKAKTYETTLK 562
 QY 361 CAAADPHCEYAKVDFEPKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPOVST 420
 DB 563 CAAADPHCEYAKVDFEPKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPOVST 622
 QY 421 PTLVEVRNLGKVGSKCKKPEAKRMPCAEDYLSVVLNOLCVLHSEKTPVSDRVTKCCTES 480
 DB 623 PTLVEVRNLGKVGSKCKKPEAKRMPCAEDYLSVVLNOLCVLHSEKTPVSDRVTKCCTES 682
 QY 481 LVNRPCFSALEVDVETVPKPEFAETFTPHADICTLSEKERQIKKOTALVELVXHKPKAT 540
 DB 683 LVNRPCFSALEVDVETVPKPEFAETFTPHADICTLSEKERQIKKOTALVELVXHKPKAT 742
 QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKGLVAASQAALGL 585
 DB 743 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKGLVAASQAALGL 787

RESULT 59

WO9724445-A1.

PD 10-JUL-1997.
XX 19-DEC-1996; 96WO-CB003164.
XX 30-DEC-1995; 95GB-00026733.
XX (DELZ) DELTA BIOTECHNOLOGY LTD.
XX Ballance DJ;
XX WPI; 1997-363680/33.
XX Serum albumin-growth hormone fusion protein - useful to treat growth hormone related diseases, e.g. Down's syndrome.
XX Example 7; Fig 11; 49pp; English.
XX Novel fusion proteins (AAW22717-20) comprise human serum albumin (HSA) joined to human growth hormone (hGH) mature polypeptide via the flexible linker (Gly-Gly-Gly-Ser)n, where n is 2, 3, 1 and 4, respectively. They are obtd. by joining HSA and hGH cDNAs (see also AAT75084 and AAT75083) via synthetic oligonucleotide linker sequences and expression in Saccharomyces cerevisiae transformants. The fusion proteins are secreted from the yeast cells and can be recovered from the culture supernatant. They show increased serum and storage stability compared with native hGH and can be used to treat growth hormone related diseases or to stimulate growth and meat production in farm animals. (Updated on 27-AUG-2003 to correct OS field.)
XX Sequence 779 AA;
SQ Query Match 99.8%; Score 584; DB 2; Length 779;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 AHKSEVAHFVKDGENFKALVLIAPVLAQYLOQCPEDHVKLVNEVTEFAKTCVADESAEN 61
DB 1 AHKSEVAHFVKDGENFKALVLIAPVLAQYLOQCPEDHVKLVNEVTEFAKTCVADESAEN 60
QY 62 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHDDNPNIPLVRPEVD 121
DB 61 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHDDNPNIPLVRPEVD 120
QY 122 VMCTAFHDEETFLKYLVEIARRHPYFVAPPELLFFAKYKAAFTCCQAAADKACLLPK 181
DB 121 VMCTAFHDEETFLKYLVEIARRHPYFVAPPELLFFAKYKAAFTCCQAAADKACLLPK 180
QY 182 LDELREDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLTKV 241
DB 181 LDELREDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLTKV 240
QY 242 HTCCCHGDLLECCADDDRADLAKYICENODSISKLKCECKPILKSHCIAEVENDEMPAD 301
DB 241 HTCCCHGDLLECCADDDRADLAKYICENODSISKLKCECKPILKSHCIAEVENDEMPAD 300
QY 302 LPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKYVETTLKCC 361
DB 301 LPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKYVETTLKCC 360
QY 362 AAADPHCYAKVDFEFPKPIVEEONLIKQNCLEFQGLGEYKQNALVRYTKKVPQVSTP 421
DB 361 AAADPHCYAKVDFEFPKPIVEEONLIKQNCLEFQGLGEYKQNALVRYTKKVPQVSTP 420
QY 422 TLVEVSRLGKVGSKCKKHPKARMPCAEDYLSVLNQLCVLHKETPPVSDRVTKCTESL 481
DB 421 TLVEVSRLGKVGSKCKKHPKARMPCAEDYLSVLNQLCVLHKETPPVSDRVTKCTESL 480
QY 482 VNRPCPSALEVDITYVPKFNATETFFHADICTLSEKROQIKQTALVELVKKHPKATK 541
DB 481 VNRPCPSALEVDITYVPKFNATETFFHADICTLSEKROQIKQTALVELVKKHPKATK 540
QY 542 BQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585

DB 541 BQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 584
RESULT 61
AAW22717
ID AAW22717 standard; protein; 784 AA.
XX AAW22717;
XX 27-AUG-2003 (revised)
DT 08-OCT-1997 (first entry)
XX Human serum albumin-human growth hormone fusion protein (n=2).
XX Growth hormone; somatostatin; hGH; human serum albumin; HSA;
KW fusion protein; protein secretion; yeast; Saccharomyces cerevisiae;
KW feed additive; Down's syndrome.
XX Homo sapiens.
OS Chimeric.
OS Chimeric.
OS Synthetic.
XX Key Location/Qualifiers
FT Protein 1..584
FT Peptide /label= HSA
FT Peptide 585..594
FT Protein /label= Linker
FT Protein 595..784
FT Protein /label= hGH
XX W09724445-A1.
XX 10-JUL-1997.
XX 19-DEC-1996; 96WO-GB003164.
XX 30-DEC-1995; 95GB-00026733.
XX (DELZ) DELTA BIOTECHNOLOGY LTD.
XX Ballance DJ;
XX WPI; 1997-363680/33.
XX Serum albumin-growth hormone fusion protein - useful to treat growth hormone related diseases, e.g. Down's syndrome.
XX Example 7; Fig 11; 49pp; English.
XX Novel fusion proteins (AAW22717-20) comprise human serum albumin (HSA) joined to human growth hormone (hGH) mature polypeptide via the flexible linker (Gly-Gly-Gly-Ser)n, where n is 2, 3, 1 and 4, respectively. They are obtd. by joining HSA and hGH cDNAs (see also AAT75084 and AAT75083) via synthetic oligonucleotide linker sequences and expression in Saccharomyces cerevisiae transformants. The fusion proteins are secreted from the yeast cells and can be recovered from the culture supernatant. They show increased serum and storage stability compared with native hGH and can be used to treat growth hormone related diseases or to stimulate growth and meat production in farm animals. (Updated on 27-AUG-2003 to correct OS field.)
XX Sequence 784 AA;
SQ Query Match 99.8%; Score 584; DB 2; Length 784;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 AHKSEVAHFVKDGENFKALVLIAPVLAQYLOQCPEDHVKLVNEVTEFAKTCVADESAEN 61
DB 1 AHKSEVAHFVKDGENFKALVLIAPVLAQYLOQCPEDHVKLVNEVTEFAKTCVADESAEN 60
QY 62 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHDDNPNIPLVRPEVD 121

```
Db 61 CDKSLHTLFGDKLCTVATLRETYGEWADCCAKQEPERNECFLOHKDDNENLRLVRPEVD 120
Qy 122 VMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQADKAACLTPK 181
Db 121 VMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQADKAACLTPK 180
Qy 182 LDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAVSKLVTDLTKV 241
Db 181 LDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAVSKLVTDLTKV 240
Qy 242 HTECHGDLLECCADRADLAKYICENQDSISSKLECCCKPPLLEKSHCIAEVENDEMPAD 301
Db 241 HTECHGDLLECCADRADLAKYICENQDSISSKLECCCKPPLLEKSHCIAEVENDEMPAD 300
Qy 302 LPSLAADFVSKQVCKNYABAKDVFGLMFLEYARRHPDYSVLLRLAKTYETTLKCC 361
Db 301 LPSLAADFVSKQVCKNYABAKDVFGLMFLEYARRHPDYSVLLRLAKTYETTLKCC 360
Qy 362 AAADPHECYAKVDFEFPKPLVEEPQNLKQNCSELFQELGEYKFNALLVRYTKVPQVSTP 421
Db 361 AAADPHECYAKVDFEFPKPLVEEPQNLKQNCSELFQELGEYKFNALLVRYTKVPQVSTP 420
Qy 422 TLVEVSRNLGKVGSKCKHPKAKMPCAEEDYLSVNLNQLCVLHEKTPVSDRVTKCTTESL 481
Db 421 TLVEVSRNLGKVGSKCKHPKAKMPCAEEDYLSVNLNQLCVLHEKTPVSDRVTKCTTESL 480
Qy 482 VNRPCFSALAEVDETYVPKFNAAETFFHADICTLSEKERQIKKQALVELVGHKPKATK 541
Db 481 VNRPCFSALAEVDETYVPKFNAAETFFHADICTLSEKERQIKKQALVELVGHKPKATK 540
Qy 542 EQLKXVMDDDFAAFVEKCKCADDKETCTFAEEGKKLVAASQAALGL 585
Db 541 EQLKXVMDDDFAAFVEKCKCADDKETCTFAEEGKKLVAASQAALGL 584
```

RESULT 62

AAW22718
ID AAW22718 standard; protein; 789 AA.

XX AC AAW22718;

DT 27-AUG-2003 (revised)

DT 08-OCT-1997 (first entry)

XX DE Human serum albumin-human growth hormone fusion protein (n=3).

KW Growth hormone; somatostatin; hGH; human serum albumin; HSA;
KW fusion protein; protein secretion; yeast; Saccharomyces cerevisiae;
KW feed additive; Down's syndrome.

XX OS Homo sapiens.
XX OS Chimaeeric.
XX OS Chimeric.
XX OS Synthetic.

EH Key Location/Qualifiers
FT Protein 1..584
FT Peptide /label= HSA
FT Peptide 585..599
FT Protein /label= Linker
FT Protein 600..789
FT Protein /label= hGH

XX WO9724445-A1.

XX 10-JUL-1997.

XX 19-DEC-1996; 96WO-GB003164.

XX 30-DEC-1995; 95GB-00026733.

XX (DELZ) DELTA BIOTECHNOLOGY LTD.

XX Ballance DJ;
XX WPI; 1997-363680/33.
XX Serum albumin-growth hormone fusion protein - useful to treat growth
XX hormone related diseases, e.g. Down's syndrome.

XX Example 7; Fig 11; 49pp; English.

XX Novel fusion proteins (AAW22717-20) comprise human serum albumin (HSA)
XX joined to human growth hormone (hGH) mature polypeptide via the flexible
XX linker (Gly-Gly-Gly-Ser)n, where n is 2, 3, 1 and 4, respectively.
XX They are obd. by joining HSA and hGH cDNAs (see also AAT75084 and
XX AAT75083) via synthetic oligonucleotide linker sequences and expression
XX in Saccharomyces cerevisiae transformants. The fusion proteins are
XX secreted from the yeast cells and can be recovered from the culture
XX supernatant. They show increased serum and storage stability compared
XX with native hGH and can be used to treat growth hormone related diseases
XX or to stimulate growth and meat production in farm animals. (Updated on
XX 27-AUG-2003 to correct OS field.)

XX Sequence 789 AA;

Query Match 99.8%; Score 584; DB 2; Length 789;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 2 AHKSEVAHRFKDLGEENFKALVLIAPQYLQCCPPEDHVKLVNEVTEFAKTCVADESAN 61
Db 1 AHKSEVAHRFKDLGEENFKALVLIAPQYLQCCPPEDHVKLVNEVTEFAKTCVADESAN 60
Qy 62 CDKSLHTLFGDKLCTVATLRETYGEWADCCAKQEPERNECFLOHKDDNENLRLVRPEVD 121
Db 61 CDKSLHTLFGDKLCTVATLRETYGEWADCCAKQEPERNECFLOHKDDNENLRLVRPEVD 120
Qy 122 VMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQADKAACLTPK 181
Db 121 VMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQADKAACLTPK 180
Qy 182 LDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAVSKLVTDLTKV 241
Db 181 LDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAVSKLVTDLTKV 240
Qy 242 HTECHGDLLECCADRADLAKYICENQDSISSKLECCCKPPLLEKSHCIAEVENDEMPAD 301
Db 241 HTECHGDLLECCADRADLAKYICENQDSISSKLECCCKPPLLEKSHCIAEVENDEMPAD 300
Qy 302 LPSLAADFVSKQVCKNYABAKDVFGLMFLEYARRHPDYSVLLRLAKTYETTLKCC 361
Db 301 LPSLAADFVSKQVCKNYABAKDVFGLMFLEYARRHPDYSVLLRLAKTYETTLKCC 360
Qy 362 AAADPHECYAKVDFEFPKPLVEEPQNLKQNCSELFQELGEYKFNALLVRYTKVPQVSTP 421
Db 361 AAADPHECYAKVDFEFPKPLVEEPQNLKQNCSELFQELGEYKFNALLVRYTKVPQVSTP 420
Qy 422 TLVEVSRNLGKVGSKCKHPKAKMPCAEEDYLSVNLNQLCVLHEKTPVSDRVTKCTTESL 481
Db 421 TLVEVSRNLGKVGSKCKHPKAKMPCAEEDYLSVNLNQLCVLHEKTPVSDRVTKCTTESL 480
Qy 482 VNRPCFSALAEVDETYVPKFNAAETFFHADICTLSEKERQIKKQALVELVGHKPKATK 541
Db 481 VNRPCFSALAEVDETYVPKFNAAETFFHADICTLSEKERQIKKQALVELVGHKPKATK 540
Qy 542 EQLKXVMDDDFAAFVEKCKCADDKETCTFAEEGKKLVAASQAALGL 585
Db 541 EQLKXVMDDDFAAFVEKCKCADDKETCTFAEEGKKLVAASQAALGL 584
```

RESULT 63

AAW22720

ID AAW22720 standard; protein; 794 AA.

XX

AAW22720;
 27-AUG-2003 (revised)
 08-OCT-1997 (first entry)
 Human serum albumin-human growth hormone fusion protein (n=4).
 Growth hormone; somatostatin; hGH; human serum albumin; HSA;
 fusion protein; protein secretion; yeast; Saccharomyces cerevisiae;
 feed additive; Down's syndrome.
 Homo sapiens.
 Chimeric.
 Synthetic.
 Key Location/Qualifiers
 Protein 1..584
 /label= HSA
 Peptide 585..604
 /label= Linker
 Protein 605..794
 /label= hGH
 WO9724445-A1.
 10-JUL-1997.
 19-DEC-1996; 96WO-GB003164.
 30-DEC-1995; 95GB-00026733.
 (DELZ) DELTA BIOTECHNOLOGY LTD.
 Ballance DJ;
 WPI; 1997-363680/33.
 Serum albumin-growth hormone fusion protein - useful to treat growth hormone related diseases, e.g. Down's syndrome.
 Example 7; Fig 11; 49pp; English.
 Novel fusion proteins (AAW22717-20) comprise human serum albumin (HSA) joined to human growth hormone (hGH) mature polypeptide via the flexible linker (Gly-Gly-Gly-Ser)n, where n is 2, 3, 1 and 4, respectively. They are obtd. by joining HSA and hGH cDNAs (see also AAT75084 and AAT75083) via synthetic oligonucleotide linker sequences and expression in Saccharomyces cerevisiae transformants. The fusion proteins are secreted from the yeast cells and can be recovered from the culture supernatant. They show increased serum and storage stability compared with native hGH and can be used to treat growth hormone related diseases or to stimulate growth and meat production in farm animals. (Updated on 27-AUG-2003 to correct OS field.)
 Sequence 794 AA;
 Query Match 99.8%; Score 584; DB 2; Length 794;
 Best Local Similarity 100.0%; Pred.No. 0;
 Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 2 AHKSEVAHRKFDLGEENFKALVIAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAEN 61
 1 AHKSEVAHRKFDLGEENFKALVIAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAEN 60
 62 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEVD 121
 61 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEVD 120
 122 VMCTAFHDNEETFLKKVLYEIARRHPVYAPPELLFFAKYKAAFTCCOAAADKAACLLPK 181
 121 VMCTAFHDNEETFLKKVLYEIARRHPVYAPPELLFFAKYKAAFTCCOAAADKAACLLPK 180

182 LDELREGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAVSKVLTDLTKV 241
 181 LDELREGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAVSKVLTDLTKV 240
 242 HTECHGDLLECCADRADLAKYICENQDSISSKLEKCECKPLEKSHCHIAEVENDEMPAD 301
 241 HTECHGDLLECCADRADLAKYICENQDSISSKLEKCECKPLEKSHCHIAEVENDEMPAD 300
 302 LPSLAADFVESKDVCKNYAKADVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKCC 361
 301 LPSLAADFVESKDVCKNYAKADVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKCC 360
 362 AAADPHECYAKVDFDEPKPLVEEPPNLIKQNCLEFQOLGEYKFKQNALLVRYTKVPQVSTP 421
 361 AAADPHECYAKVDFDEPKPLVEEPPNLIKQNCLEFQOLGEYKFKQNALLVRYTKVPQVSTP 420
 422 TLVEVSRNLGKVGSKCKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTVPVSDRVTCKCTESL 481
 421 TLVEVSRNLGKVGSKCKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTVPVSDRVTCKCTESL 480
 482 VNRPPCFSALEVDETYVPKFNATETFTFHADICTLSEKEROIKKQTALVELVGHKPKATK 541
 481 VNRPPCFSALEVDETYVPKFNATETFTFHADICTLSEKEROIKKQTALVELVGHKPKATK 540
 542 EQLKAVMDDFAAVFEKCKCKADDETCFAEKGKLVAAASQAALGL 585
 541 EQLKAVMDDFAAVFEKCKCKADDETCFAEKGKLVAAASQAALGL 584
 RESULT 64
 AAR92149
 ID AAR92149 standard; protein; 754 AA.
 XX
 AC AAR92149;
 XX
 DT 16-OCT-2003 (revised)
 DT 10-SEP-1996 (first entry)
 XX
 DE HSA:Fc gamma RII fusion protein.
 XX
 KW Fc receptor; Fc gamma RII; IGG; autoimmune disease; therapy; antagonist;
 KW human serum albumin; HSA; immunoglobulin.
 XX
 OS Homo sapiens.
 OS Chimeric.
 XX
 PN WO9608512-A1.
 XX
 PD 21-MAR-1996.
 XX
 PF 15-SEP-1995; 95WO-AU000606.
 XX
 PR 16-SEP-1994; 94AU-00008232.
 PR 31-OCT-1994; 94US-00332562.
 XX
 PA (AUST-) AUSTIN RES INST CANCER & ANTI INFLAMMATO.
 XX
 PI Hogarth PM, McKenzie I, Baker RI, Hulett MD, Powell MS;
 XX
 XX WPI; 1996-179903/18.
 DR N-PSDB; AAT14530.
 XX
 XX New mutant Fc receptor polypeptide(s) - have amino acid changes to
 PT improve characteristics, e.g. half life, used partic in diagnosis or
 PT treatment of auto-immune diseases.
 XX
 PS Claim 21; Fig 12; 104pp; English.
 XX
 CC A fusion protein (AAR92149), HSA:Fc gamma RII, comprises human serum
 CC albumin (HSA) fused at its C-terminal end to the extracellular domains of
 CC Fc gamma RII. It was obtd. by transformation of Pichia pastoris cells
 CC using a gene fusion (AAT14530) obtd. by splice overlap extension PCR.
 CC The 100 kDa fusion protein retains the Fc receptor activity of Fc gamma

CCC	RII. It can be used to detect immunoglobulins (Ig), remove Ig from samples, screen for antagonists, or to treat diseases associated with excess Ig. The half-life in mice is 140 min compared to 40 min for Fc gamma RII. (Updated on 16-OCT-2003 to standardise OS field)
CCX	
XX	
SSQ	Sequence 754 AA;
	Query Match 96.4%; Score 564; DB 2; Length 754;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 DAHKSEVAHFKDLGRNFKALVLIIFAQYLQOCPEFDHVKLIVNEVFIAKTCVADESAAE 60
Dd	1 DAHKSEVAHFKDLGRNFKALVLIIFAQYLQOCPEFDHVKLIVNEVFIAKTCVADESAAE 60
QY	61 NCDKSLHTLFGDKLT VATILRETYGEMADCCAKQSPERNECPLOKHDDNPMLPLVRPEV 120
Dd	61 NCDKSLHTLFGDKLT VATILRETYGEMADCCAKQSPERNECPLOKHDDNPMLPLVRPEV 120
QY	121 DVMCTAHDNEETFLKYLYEIARRHPYFVAPELLFFAKRYKAAATECCQAADKAACLLP 180
Dd	121 DVMCTAHDNEETFLKYLYEIARRHPYFVAPELLFFAKRYKAAATECCQAADKAACLLP 180
QY	181 KLDELDEGRASSAKORLKCSAQKFGERAFKAWAVARLSQRFPKAEVSKLVTDLTK 240
Dd	181 KLDELDEGRASSAKORLKCSAQKFGERAFKAWAVARLSQRFPKAEVSKLVTDLTK 240
QY	241 VHTCCGGDLLLECADDRADI AKVICENQDSISSKLEKCECKELLESKSHCIAEVNDMPA 300
Dd	241 VHTCCGGDLLLECADDRADLA KVICENQDSISSKLEKCECKELLESKSHCIAEVNDMPA 300
QY	301 DLPSLAADFVESKDVCNKVAEKDVLGMFLPYEARRHDPYSVWLLRLAKTYETTLEKC 360
Dd	301 DLPSLAADFVESKDVCNKVA EKDVLGMFLPYEARRHDPYSVWLLRLAKTYETTLEKC 360
QY	361 CAADAPHECYAKVDFEPKPIVEBPQNLIKNCBELPOLGEYFQNALIVRYTKVPQVST 420
Dd	361 CAADAPHECYAKVDFEPKPIVEBPQNLIKNCBELPOLGEYFQNALIVRYTKVPQVST 420
QY	421 PTLVEVSRLKGVGSKCKHPEAKMPCASDYLSVLNLQCVLHEKTPVSDRVTKCCTES 480
Dd	421 PTLVEVSRLKGVGSKCKHPEAKMPCASDYLSVLNLQCVLHEKTPVSDRVTKCCTES 480
QY	481 LVNRRCPCFSALEVEDTVYPKEFNAAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540
Dd	481 LVNRRCPCFSALEVEDTVYPKEFNAAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540
QY	541 KEQLKAVMDDFAAFVEKCKCKADDK 564
Dd	541 KEQLKAVMDDFAAFVEKCKCKADDK 564
RESULT 65	
ABP98782	ID ABP98782 standard; peptide; 585 AA.
XX	AC AC
XX	ABP98782;
XX	DT DT
XX	DE DE
XX	Human serum albumin with selenocysteine at position 34.
KW	Human serum albumin; selenocysteine; immunomodulator; cytostatic;
KW	cardiant; nootropic; neuroprotective; antidiabetic; ophthalmological;
KW	selenoprotein; selenocysteine; oxidative stress; aging; inflammation;
KW	neurodegeneration; diabetes; cataract; atrophy.
XX	Homo sapiens.
OS	XX
XX	Key Location/Qualifiers
FT	Modified-site 34
FT	/note= "selenocysteine amino acid"
XX	XX

WN	W02003029469-AI.
XX	
PD	10-APR-2003.
XX	
PF	12-SEP-2002; 2002WO-JP09313.
XX	
PP	13-SEP-2001; 2001JP-00278749.
XX	
XX	(XAGA) CHEMO-SERO-THERAPEUTIC RES INST.
XX	Kaminaka S, Kaminaka K, Hirashima M, Maeda H, Nozaki C;
PFI	Takahashi K;
XX	
XX	WPI; 2003-354730/33.
XX	
PT	Novel selenocysteine-containing proteins with phospholipid peroxide-
XX	reducing activity and encoded genes, applicable in inhibition,
XX	prevention, treatment or deterioration of e.g. aging and inflammation.
XX	
PS	Claim 4; Page 31-33; 42pp; Japanese.
XX	
CC	The invention relates to a method of generating a selenoprotein by the
CC	insertion or substitution of one or more selenocysteines into the
CC	skeleton of a non-selenocysteine-containing protein. The method is
CC	especially targeted to proteins having enzyme activity. The proteins have
CC	applications as antioxidant substance in inhibition, prevention,
CC	treatment or deterioration of physiological conditions due to oxidative
CC	stress e.g. aging, inflammation, neurodegeneration, diabetes, cataract
CC	and atrophy. In an example of invention, the method is used to introduce
CC	a selenocysteine into the human serum albumin (HSA) protein. This
CC	sequence represents the mature form of an HSA protein containing a
CC	selenocysteine at amino acid position 34
XX	
SQ	Sequence 585 AA;
	Query Match 94.2%; Score 551; DB 6; Length 585;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 551; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	35 PFEDHVKLNVNEVTEFAKTCVADESACNDKSLHTLFGDKLCTVATLRITYGEMADCCAQ 94
DB	35 PFEDHVKLNVNEVTEFAKTCVADESACNDKSLHTLFGDKLCTVATLRITYGEMADCCAQ 94
QY	95 EPERNECFLOHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKLYLETARHPVFYAPEL 154
DB	95 EPERNECFLOHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKLYLETARHPVFYAPEL 154
QY	155 LFFAKRYKAATFECCQAADAAACLLPKLDLDRDEGKASSAKORLKCSAQKFGERAFKAW 214
DB	155 LFFAKRYKAATFECCQAADAAACLLPKLDLDRDEGKASSAKORLKCSAQKFGERAFKAW 214
QY	215 AVAHSORFPKAEPAEVSUKLVTLTKVHTCECHGDILLECADRADLAKYICENQDSTSSK 274
DB	215 AVAHSORFPKAEPAEVSUKLVTLTKVHTCECHGDILLECADRADLAKYICENQDSTSSK 274
QY	275 LKECCEKPLEBKSHCIAEVENDMPADLPSLAADFVESKDVCNKYAEAKDVFLGMFLVEY 334
DB	275 LKECCEKPLEBKSHCIAEVENDMPADLPSLAADFVESKDVCNKYAEAKDVFLGMFLVEY 334
QY	335 ARRPDPYSVWLILLAKTYETTILEKCAAADPHCEYAKVPDFEPKPLVBEPQNLIKONCEL 394
DB	335 ARRPDPYSVWLILLAKTYETTILEKCAAADPHCEYAKVPDFEPKPLVBEPQNLIKONCEL 394
QY	395 FPELGXEYFQNALIVRTTKVPQVSTPLLVEVSRNLGKVGSCKCKHPRAKNPCDAEDYLS 454
DB	395 FPELGXEYFQNALIVRTTKVPQVSTPLLVEVSRNLGKVGSCKCKHPRAKNPCDAEDYLS 454
QY	455 VVLNQLCVLEHEKTVPVSDRVTKCCTESILVNRPCFSALEVDETYVPKFEVNAETFTFHADIC 514
DB	455 VVLNQLCVLEHEKTVPVSDRVTKCCTESILVNRPCFSALEVDETYVPKFEVNAETFTFHADIC 514
QY	515 TLSEKERQIKQTALVELVGHKPKATEQLKAYMDDFAAFEVKCCCKADDKETCFEAEGKK 574

Db 515 TLSEKERQIKQTALVELVGHKPKATKEQLKAVMDFAAFVEKCKADKDKETCFABEGKK 574
 QY 575 LVAASQAALGL 585
 Db 575 LVAASQAALGL 585

RESULT 66
 ABP98783
 ID ABP98783 standard; peptide; 620 AA.
 XX
 AC ABP98783;
 DT 25-JUL-2003 (first entry)
 DE Full length human serum albumin with selenocysteine at position 58.
 XX
 KW Human serum albumin; selenocysteine; immunomodulator; cytostatic;
 KW cardiant; neurotropic; neuroprotective; antidiabetic; ophthalmological;
 KW selenoprotein; selenocysteine; oxidative stress; aging; inflammation;
 KW neurodegeneration; diabetes; cataract; atrophy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 58
 FT /label= OTHER
 FT /note= "OTHER = selenocysteine"
 XX
 PN WO2003029459-A1.
 XX
 XX 10-APR-2003.
 XX
 XX 12-SEP-2002; 2002WO-JP009313.
 XX
 XX 13-SEP-2001; 2001JP-00278749.
 XX
 XX (KAGA) CHEMO-SERO-THERAPEUTIC RES INST.
 XX
 XX Kaminaka S, Kaminaka K, Hiraishima M, Maeda H, Nozaki C;
 XX Takahashi K;
 XX
 XX WPI; 2003-354730/33.
 XX
 XX N-PSDB; ABZ80909.
 XX
 XX Novel selenocysteine-containing proteins with phospholipid peroxide-
 XX reducing activity and encoded genes, applicable in inhibition,
 XX prevention, treatment or deterioration of e.g. aging and inflammation.
 XX
 XX Example 1; Page 34-37; 42pp; Japanese.
 XX
 XX The invention relates to a method of generating a selenoprotein by the
 XX insertion or substitution of one or more selenocysteines into the
 XX skeleton of a non-selenocysteine-containing protein. The method is
 XX especially targeted to proteins having enzyme activity. The proteins have
 XX applications as antioxidants of physiological conditions due to oxidative
 XX treatment or deterioration of physiological conditions due to oxidative
 XX stress e.g. aging, inflammation, neurodegeneration, diabetes, cataract
 XX and atrophy. In an example of invention the method is used to introduce
 XX a selenocysteine into the human serum albumin (HSA) protein. This
 XX sequence represents the full length form of an HSA protein containing a
 XX selenocysteine at amino acid position 58
 XX
 XX Sequence 620 AA;
 SQ

Query Match 94.2%; Score 551; DB 6; Length 620;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 551; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 35 PFEDHVLVNEVTEFAKTCVADESAENCKSLHTLFGDKLCTVATLRETYGEMADCCAKQ 94
 Db 59 PFEDHVLVNEVTEFAKTCVADESAENCKSLHTLFGDKLCTVATLRETYGEMADCCAKQ 118

QY 95 EPERNECFLOHQDDNPNLRLVRPEVDVNMCTAFHDNEETFLKKYLVEIARRHFFYAPEL 154
 Db 119 EPERNECFLOHQDDNPNLRLVRPEVDVNMCTAFHDNEETFLKKYLVEIARRHFFYAPEL 178
 QY 155 LFFAKRYKAAFTTECCQAADKAAACLLPKLDELDEGKASAKORLKASLOKFGGERAFKAW 214
 Db 179 LFFAKRYKAAFTTECCQAADKAAACLLPKLDELDEGKASAKORLKASLOKFGGERAFKAW 238
 QY 215 AVARLSORFPKAEFAEVSCLVDTLTKVHTCECHGDLLECADRADLAKYICENQDSISSK 274
 Db 239 AVARLSORFPKAEFAEVSCLVDTLTKVHTCECHGDLLECADRADLAKYICENQDSISSK 298
 QY 275 LKECCEKPLLEKSHCIAEVENDEMPADLPFLAADFVESKDVCKNYAAKDVFLGMFLYBY 334
 Db 299 LKECCEKPLLEKSHCIAEVENDEMPADLPFLAADFVESKDVCKNYAAKDVFLGMFLYBY 358
 QY 335 ARRPDPVSULLLRLAKTYETTLKCCAAADPHECYAKVDFEFKPLVERPQNLKQNCSEL 394
 Db 359 ARRPDPVSULLLRLAKTYETTLKCCAAADPHECYAKVDFEFKPLVERPQNLKQNCSEL 418
 QY 395 FEQLGEYKFNQALLVRYTKKVPQVSTPLVEVSRNLGKVGSKCKHPKAPKAPCAEDYLS 454
 Db 419 FEQLGEYKFNQALLVRYTKKVPQVSTPLVEVSRNLGKVGSKCKHPKAPKAPCAEDYLS 478
 QY 455 VTINQLCVLHEKTPVSDRVTKCTESLVNRRPCFSALEVDVTVPKFNAETFTFHADIC 514
 Db 479 VTINQLCVLHEKTPVSDRVTKCTESLVNRRPCFSALEVDVTVPKFNAETFTFHADIC 538
 QY 515 TLSEKERQIKQTALVELVGHKPKATKEQLKAVMDFAAFVEKCKADKDKETCFABEGKK 574
 Db 539 TLSEKERQIKQTALVELVGHKPKATKEQLKAVMDFAAFVEKCKADKDKETCFABEGKK 598
 QY 575 LVAASQAALGL 585
 Db 599 LVAASQAALGL 609

RESULT 67
 ADD06563
 ID ADD06563 standard; protein; 652 AA.
 XX
 AC ADD06563;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE Human Ckb1-HAS fusion protein construct secreted protein SEQ ID NO:99.
 XX
 KW human; chemokine betal; Ckb1; anti-HIV; neuroprotective; antithyroid;
 KW antiarthritic; antirheumatic; immunosuppressive; nootropic;
 KW antiinflammatory; antiasthmatic; antiallergic; osteopathia;
 KW nephrotrophic; tuberculostatic; virucide; antiatherosclerotic;
 KW antimicrobial; infection; HIV; immune disorder; haematopoietic disorder;
 KW autoimmune disorder; multiple sclerosis; Grave's disease; arthritis;
 KW rheumatoid arthritis; transplant rejection; neurodegenerative disorder;
 KW Alzheimer's disease; inflammatory disease; asthma; allergic disorder;
 KW inflammatory bowel disease; osteoarthritis; colitis;
 KW inflammatory kidney disease; glomerulonephritis; infectious disease;
 KW tuberculosis; hepatitis infection; herpes viral infection;
 KW viral infection; proliferative disorder; atherosclerosis;
 KW human serum albumin; HSA.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX WO200297038-A2.
 XX
 XX 05-DEC-2002.
 XX
 XX 24-MAY-2002; 2002WO-US016525.
 XX
 XX 25-MAY-2001; 2001US-0293212P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX PA

XX FI Bell A, Ruben SM;
 XX WPI; 2003-140456/13.
 DR N-PSDB; ADD06561.
 XX Novel human chemokine betal protein comprising deletion in amino acids
 PT from amino and/or carboxy terminus, and is a fusion protein further
 PT comprising human serum albumin, is useful for treating multiple
 PT sclerosis, asthma.
 XX Example 1; SEQ ID NO 99; 423pp; English.
 PS The present invention describes a human chemokine betal (Ckb1) protein
 CC (I) comprising a deletion in amino acid residues from the amino terminus
 CC and/or carboxy terminus of the 93 residue amino acid sequence (SI, see
 CC ADD06466). (I) has anti-HIV, neuroprotective, antithyroid, antiarthritic,
 CC antirheumatic, immunosuppressive, neutrotropic, antinflammatory,
 CC antiasthmatic, antiallergic, osteopathic, nephrotrophic, tuberculostatic,
 CC virucide, antiatherosclerotic and antimicrobial activities. (I) is useful
 CC for preventing infection, preferably viral (human immunodeficiency virus
 CC (HIV)) infection, in a cell, by contacting the cell with (I). (I) is also
 CC useful for treating a disease, such as HIV infection or immune disorders,
 CC hematopoietic disorders, autoimmune disorders, multiple sclerosis,
 CC Grave's disease, arthritis, rheumatoid arthritis, transplant rejection,
 CC neurodegenerative disorders, Alzheimer's disease, inflammatory disease,
 CC asthma, allergic disorders, inflammatory bowel disease, osteoarthritis,
 CC colitis, inflammatory kidney diseases, glomerulonephritis, infection,
 CC disease, tuberculosis, hepatitis infections, herpes viral infection,
 CC viral infection, proliferative disorders or atherosclerosis, in an
 CC individual. (I) inhibits or abolishes the ability of HIV to bind to,
 CC enter into/fuse with (infect), and/or replicate in CCR5 expressing cells.
 CC (I) also acts a CCR5 agonists or antagonists, stimulate chemotaxis of
 CC CCR5-expressing cells, inhibit CCR5 ligand binding to a CCR5 molecule, or
 CC upregulate or downregulate CCR5 expression. (I) is useful as an
 CC immunological probe for the differential identification of the tissues or
 CC cell-types. (I)-human serum albumin (HSA) fusion proteins are useful for
 CC diagnosing, treating and preventing various disorders in mammals,
 CC preferably in humans. (I)-HSA fusion proteins are also useful as
 CC molecular weight markers on sodium dodecyl sulfate polyacrylamide gel
 CC electrophoresis techniques, for raising antibodies, and to test the
 CC biological activities of the Ckb1 protein. (I)-HSA fusion proteins are
 CC useful for screening for molecules that bind to the Ckb1 protein portion
 CC of the fusion protein. The present sequence is used in the
 CC exemplification of the present invention.
 XX Sequence 652 AA;
 SQ

Query Watch 87.2%; Score 510; DB 7; Length 652;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAKSEVAHRFDLGBENFKALVLIAFAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60
 DB 67 DAKSEVAHRFDLGBENFKALVLIAFAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 126
 QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNECFLOHKDNNPLVLVPEV 120
 DB 127 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNECFLOHKDNNPLVLVPEV 186
 QY 121 DVNCTAFHDNETHFKKLYEIAHRHPYFAPELLFFAKRYKAAPTECCQAADKAACLLP 180
 DB 187 DVNCTAFHDNETHFKKLYEIAHRHPYFAPELLFFAKRYKAAPTECCQAADKAACLLP 246
 QY 181 KLDELDEGKASAKQRLKASLOKGFERAFKAVAVARLSORFPKAEFAVSKLVTDLTK 240
 DB 247 KLDELDEGKASAKQRLKASLOKGFERAFKAVAVARLSORFPKAEFAVSKLVTDLTK 306
 QY 241 VHTCCGHLLECADRADLAKYICENQDSISSKLKECCERPLLEKSKCIAEVENDEMPA 300
 DB 307 VHTCCGHLLECADRADLAKYICENQDSISSKLKECCERPLLEKSKCIAEVENDEMPA 366
 QY 301 DLPSLAADFVESKDVCKNVAEKDVLGMLFYEARRHEDYSVLLRLAKYETTTLEKC 360

DB 367 DLPSLAADFVESKDVCKNVAEKDVLGMLFYEARRHEDYSVLLRLAKYETTTLEKC 426
 QY 361 CAADAPHECYAKVDFEFKPLVPEPQNLIKQNCLEFQOLGEYKFNALLVRYTKVPQVST 420
 DB 427 CAADAPHECYAKVDFEFKPLVPEPQNLIKQNCLEFQOLGEYKFNALLVRYTKVPQVST 486
 QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 DB 487 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 546
 QY 481 LVNRRPCFSALEVDETYVPKPEFNAETFTTH 510
 DB 547 LVNRRPCFSALEVDETYVPKPEFNAETFTTH 576
 RESULT 68
 ADD06569
 ID ADD06569 standard; protein; 652 AA.
 XX AC ADD06569;
 XX DT 01-JAN-2004 (first entry)
 DE Human Ckb1-HAS fusion protein construct secreted protein SEQ ID NO:105.
 DE human; chemokine betal; Ckb1; anti-HIV; neuroprotective; antithyroid;
 KW antiarthritic; antirheumatic; immunosuppressive; neutrotropic;
 KW antinflammatory; antiasthmatic; antiallergic; osteopathic;
 KW nephrotrophic; tuberculostatic; virucide; antiatherosclerotic;
 KW antimicrobial; infection; HIV; immune disorder; haematopoietic disorder;
 KW autoimmune disorder; multiple sclerosis; Grave's disease; arthritis;
 KW rheumatoid arthritis; transplant rejection; neurodegenerative disorder;
 KW Alzheimer's disease; inflammatory disease; asthma; allergic disorder;
 KW inflammatory bowel disease; osteoarthritis; colitis;
 KW inflammatory kidney disease; glomerulonephritis; infection;
 KW tuberculosis; hepatitis infection; herpes viral infection;
 KW viral infection; proliferative disorder; atherosclerosis;
 KW human serum albumin; HSA.
 OS Synthetic.
 OS Homo sapiens.
 XX WO200297038-A2.
 XX 05-DEC-2002.
 XX 24-MAY-2002; 2002WO-US016525.
 XX 25-MAY-2001; 2001US-0293212P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Bell A, Ruben SM;
 XX WPI; 2003-140456/13.
 DR N-PSDB; ADD06567.
 PT Novel human chemokine betal protein comprising deletion in amino acids
 PT from amino and/or carboxy terminus, and is a fusion protein further
 PT comprising human serum albumin, is useful for treating multiple
 PT sclerosis, asthma.
 XX Example 1; SEQ ID NO 105; 423pp; English.
 XX The present invention describes a human chemokine betal (Ckb1) protein
 CC (I) comprising a deletion in amino acid residues from the amino terminus
 CC and/or carboxy terminus of the 93 residue amino acid sequence (SI, see
 CC ADD06466). (I) has anti-HIV, neuroprotective, antithyroid, antiarthritic,
 CC antirheumatic, immunosuppressive, neutrotropic, antinflammatory,
 CC antiasthmatic, antiallergic, osteopathic, nephrotrophic, tuberculostatic,
 CC virucide, antiatherosclerotic and antimicrobial activities. (I) is useful
 CC for preventing infection, preferably viral (human immunodeficiency virus

CC	(HIV) infection, in a cell, by contacting the cell with (I). (I) is also
CC	useful for treating a disease, such as HIV infection or immune disorders,
CC	haematopoietic disorders, autoimmune disorders, multiple sclerosis,
CC	Grave's disease, arthritis, rheumatoid arthritis, transplant rejection,
CC	neurodegenerative disorders, Alzheimer's disease, inflammatory disease,
CC	asthma, allergic disorders, inflammatory bowel disease, osteoarthritis,
CC	colitis, inflammatory kidney diseases, glomerulonephritis, infectious
CC	disease, tuberculosis, hepatitis infections, herpes viral infection,
CC	viral infection, proliferative disorders or atherosclerosis, in an
CC	individual. (I) inhibits or abolishes the ability of HIV to bind to,
CC	enter into/fuse with (infect), and/or replicate in CCR5-expressing cells.
CC	(I) also acts as CCR5 agonists or antagonists, stimulate chemotaxis of
CC	CCR5-expressing cells, inhibit CCR5 ligand binding to a CCR5 molecule, or
CC	upregulate or downregulate CCR5 expression. (I) is useful as an
CC	immunological probe for the differential identification of the tissues or
CC	cell-types. (I)-human serum albumin (HSA) fusion proteins are useful for
CC	diagnosing, treating and preventing various disorders in mammals,
CC	preferably in humans. (I)-HSA fusion proteins are also useful as
CC	molecular weight markers on sodium dodecyl sulfate polyacrylamide gel
CC	electrophoresis techniques, for raising antibodies, and to test the
CC	biological activities of the Ckbl protein. (I)-HSA fusion proteins are
CC	useful for screening for molecules that bind to the Ckbl protein portion
CC	of the fusion protein. The present sequence is used in the
CC	exemplification of the present invention.
XX	
XX	Sequence 652 AA;
SQ	
Query Match	87.2%; Score 510; DB 7; Length 652;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 510; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 DAKSEVAHRFKDGEENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB	67 DAKSEVAHRFKDGEENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 126
QY	61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKPEPERNECFLOHKDDNPNLRLVPEV 120
DB	127 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKPEPERNECFLOHKDDNPNLRLVPEV 186
QY	121 DVMCTAFHDNEETFLKLYLIEARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB	187 DVMCTAFHDNEETFLKLYLIEARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 246
QY	181 KLDELDEGKASSAKQRLKASLQKGFGRFAKAWAVARLSQFFKABFAEVSKLVTDLTK 240
DB	247 KLDELDEGKASSAKQRLKASLQKGFGRFAKAWAVARLSQFFKABFAEVSKLVTDLTK 306
QY	241 VHTCCCHGDLLECCADRDADLAKYICENQDSISSKLKECCFKPFLKSKHICIAEVENDEMPA 300
DB	307 VHTCCCHGDLLECCADRDADLAKYICENQDSISSKLKECCFKPFLKSKHICIAEVENDEMPA 366
QY	301 DPLSLAADFVSKDVCNVAEAKDVLGMPLYEYARRHPDYSVLLRLAKTYETTTLEKC 360
DB	367 DPLSLAADFVSKDVCNVAEAKDVLGMPLYEYARRHPDYSVLLRLAKTYETTTLEKC 426
QY	361 CAAADPHECVAKYVDFEFLKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKKVPQVST 420
DB	427 CAAADPHECVAKYVDFEFLKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKKVPQVST 486
QY	421 PTLVEVSRNLGKVGSKCKKHPRKAMPQCAEDYLSVVLNQLCVLHKTPVSRVTKCCTES 480
DB	487 PTLVEVSRNLGKVGSKCKKHPRKAMPQCAEDYLSVVLNQLCVLHKTPVSRVTKCCTES 546
QY	481 LVNRRPFCFSALEVDENVYVPKEFNAETFTFH 510
DB	547 LVNRRPFCFSALEVDENVYVPKEFNAETFTFH 576
RESULT 69	
ADD06560	
ID	ADD06560 standard; protein; 652 AA.
XX	
AC	ADD06560;

XX	01-JAN-2004 (first entry)
DT	Human Ckbl-HAS fusion protein construct secreted protein SEQ ID NO:96.
XX	human; chemokine betal; Ckbl; anti-HIV; neuroprotective; antithyroid;
XX	antiarthritic; antirheumatic; immunosuppressive; nootropic;
DE	antiinflammatory; antiasthmatic; anti-allergic; osteopathic;
XX	nephrotrophic; tuberculostatic; virucide; antiatherosclerotic;
XX	antimicrobial; infection; HIV; immune disorder; haematopoietic disorder;
XX	autoimmune disorder; multiple sclerosis; Grave's disease; arthritis;
XX	rheumatoid arthritis; transplant rejection; neurodegenerative disorder;
XX	Alzheimer's disease; inflammatory disease; asthma; allergic disorder;
XX	inflammatory bowel disease; osteoarthritis; colitis;
XX	inflammatory kidney disease; glomerulonephritis; infectious disease;
XX	tuberculosis; hepatitis infection; herpes viral infection;
XX	viral infection; proliferative disorder; atherosclerosis;
XX	human serum albumin; HSA.
OS	Synthetic.
OS	Homo sapiens.
XX	WO200297038-A2.
PN	05-DEC-2002.
XX	24-MAY-2002; 2002WO-US016525.
PF	25-MAY-2001; 2001US-0293212P.
XX	(HUMA-) HUMAN GENOME SCI INC.
PA	Bell A, Ruben SM;
XX	WPI; 2003-140456/13.
PI	N-PSDB; ADD06558.
XX	Novel human chemokine betal protein comprising deletion in amino acids
XX	from amino and/or carboxy terminus, and is a fusion protein further
XX	comprising human serum albumin, is useful for treating multiple
XX	sclerosis, asthma.
XX	Example 1; SEQ ID NO 96; 423pp; English.
XX	The present invention describes a human chemokine betal (Ckbl) protein
XX	(I) comprising a deletion in amino acid residues from the amino terminus
XX	and/or carboxy terminus of the 93 residue amino acid sequence (SI, see
XX	ADD06466). (I) has anti-HIV, neuroprotective, antithyroid, antiarthritic,
XX	antirheumatic, immunosuppressive, nootropic, antiinflammatory,
XX	antiasthmatic, anti-allergic, osteopathic, nephrotrophic, tuberculostatic,
XX	virucide, antiatherosclerotic and antimicrobial activities. (I) is useful
XX	for preventing infection, preferably viral (human immunodeficiency virus
XX	(HIV)) infection, in a cell, by contacting the cell with (I). (I) is also
XX	useful for treating a disease, such as HIV infection or immune disorders,
XX	haematopoietic disorders, autoimmune disorders, multiple sclerosis,
XX	Grave's disease, arthritis, rheumatoid arthritis, transplant rejection,
XX	neurodegenerative disorders, Alzheimer's disease, inflammatory disease,
XX	asthma, allergic disorders, inflammatory bowel disease, osteoarthritis,
XX	colitis, inflammatory kidney diseases, glomerulonephritis, infectious
XX	disease, tuberculosis, hepatitis infections, herpes viral infection,
XX	viral infection, proliferative disorders or atherosclerosis, in an
XX	individual. (I) inhibits or abolishes the ability of HIV to bind to,
XX	enter into/fuse with (infect), and/or replicate in CCR5-expressing cells.
XX	(I) also acts as CCR5 agonists or antagonists, stimulate chemotaxis of
XX	CCR5-expressing cells, inhibit CCR5 ligand binding to a CCR5 molecule, or
XX	upregulate or downregulate CCR5 expression. (I) is useful as an
XX	immunological probe for the differential identification of the tissues or
XX	cell-types. (I)-human serum albumin (HSA) fusion proteins are useful for
XX	diagnosing, treating and preventing various disorders in mammals,
XX	preferably in humans. (I)-HSA fusion proteins are also useful as
XX	molecular weight markers on sodium dodecyl sulfate polyacrylamide gel
XX	electrophoresis techniques, for raising antibodies, and to test the
XX	biological activities of the Ckbl protein. (I)-HSA fusion proteins are

CC useful for screening for molecules that bind to the Cxbl protein portion
CC of the fusion protein. The present sequence is used in the
CC exemplification of the present invention.

XX
SQ Sequence 652 AA;

Query Match 87.2%; Score 510; DB 7; Length 652;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGEENFKALVLIAPQYLOQCPEDEHVKLVNEVTEFAKTCVADESAAE 60
DB 67 DAHSEVAHFRKDLGEENFKALVLIAPQYLOQCPEDEHVKLVNEVTEFAKTCVADESAAE 126

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLVLRPEV 120
DB 127 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLVLRPEV 186

QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFFAKRYKAAFTCCQAADKAAACLLP 180
DB 187 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFFAKRYKAAFTCCQAADKAAACLLP 246

QY 181 KLDELREGKASAKQRLKASLOKGERAFKAWAVARLSQRPFAEFAVSKLVTLDTK 240
DB 247 KLDELREGKASAKQRLKASLOKGERAFKAWAVARLSQRPFAEFAVSKLVTLDTK 306

QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCECEKPLEKSHCIAEVENDEMPA 300
DB 307 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCECEKPLEKSHCIAEVENDEMPA 366

QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGFLFYEARHPDVSIVLLKATYTTTLEKC 360
DB 367 DLPSLAADFVESKDVCKNVAEAKDVLGFLFYEARHPDVSIVLLKATYTTTLEKC 426

QY 361 CAADPHCEYAKVDFEFKPLVEBPQNLKQNCLEPEQLGEYKFNALLVRYTKVPQVST 420
DB 427 CAADPHCEYAKVDFEFKPLVEBPQNLKQNCLEPEQLGEYKFNALLVRYTKVPQVST 486

QY 421 PTLVENSRLGKVGSKCKRHPKAPCAEDVLSVLNQLCVLHKTPSVDRVTKCTTES 480
DB 487 PTLVENSRLGKVGSKCKRHPKAPCAEDVLSVLNQLCVLHKTPSVDRVTKCTTES 546

QY 481 LVNRPCFSALEVDETVVPKEFNAETFTFH 510
DB 547 LVNRPCFSALEVDETVVPKEFNAETFTFH 576

RESULT 70
ADD06557
ID ADD06557 standard; protein: 660 AA.
AC ADD06557;
XX
XX
DT 01-JAN-2004 (first entry)
XX
DE Human Cxbl-HAS fusion protein construct secreted protein SEQ ID NO:93.
XX
KW human; chemokine betal; Cxbl; anti-HIV; neuroprotective; antithyroid;
KW antiarthritic; antirheumatic; immunosuppressive; nontropic;
KW antiinflammatory; antiasthmatic; antiallergic; osteopathic;
KW nephrotrophic; tuberculostatic; virucide; antiatherosclerotic;
KW antimicrobial; infection; HIV; immune disorder; haematopoietic disorder;
KW autoimmune disorder; multiple sclerosis; Grave's disease; arthritis;
KW rheumatoid arthritis; transplant rejection; neurodegenerative disorder;
KW inflammatory bowel disease; inflammatory disease; asthma; allergic disorder;
KW inflammatory kidney disease; osteoarthritis; colitis;
KW inflammatory disease; glomerulonephritis; infectious disease;
KW tuberculosi; hepatitis infection; herpes viral infection;
KW viral infection; proliferative disorder; atherosclerosis;
KW human serum albumin; HSA.
OS Synthetic.
OS Homo sapiens.

WO200297038-A2.

05-DEC-2002.

24-MAY-2002; 2002WO-US016525.

25-MAY-2001; 2001US-0293212P.

(HUMA-) HUMAN GENOME SCI INC.

Bell A, Ruben SM;

WPI; 2003-140456/13.

N-PSDB; ADD06555.

Novel human chemokine betal protein comprising deletion in amino acids from amino and/or carboxy terminus, and is a fusion protein further comprising human serum albumin, is useful for treating multiple sclerosis, asthma.

Example 1; SEQ ID NO 93; 423pp; English.

The present invention describes a human chemokine betal (Cxbl) protein (I) comprising a deletion in amino acid residues from the amino terminus and/or carboxy terminus of the 93 residue amino acid sequence (S1, see ADD06466). (I) has anti-HIV, neuroprotective, antithyroid, antiarthritic, antirheumatic, immunosuppressive, nontropic, antiinflammatory, antiasthmatic, antiallergic, osteopathic, nephrotrophic, tuberculostatic, virucide, antiatherosclerotic and antimicrobial activities. (I) is useful for preventing infection, preferably viral (human immunodeficiency virus (HIV)) infection, in a cell, by contacting the cell with (I). (I) is also useful for treating a disease, such as HIV infection or immune disorders, haematopoietic disorders, autoimmune disorders, multiple sclerosis, Grave's disease, arthritis, rheumatoid arthritis, transplant rejection, neurodegenerative disorders, Alzheimer's disease, inflammatory disease, asthma, allergic disorders, inflammatory bowel disease, osteoarthritis, colitis, inflammatory kidney diseases, glomerulonephritis, infectious disease, tuberculosi, hepatitis infections, herpes viral infection, in an individual. (I) inhibits or abolishes the ability of HIV to bind to, enter into/fuse with (infect), and/or replicate in CCR5 expressing cells. (I) also acts as CCR5 agonists or antagonists, stimulate chemotaxis of CCR5-expressing cells inhibit CCR5 ligand binding to a CCR5 molecule, or upregulate or downregulate CCR5 expression. (I) is useful as an immunological probe for the differential identification of the tissues or cell-types. (I)-human serum albumin (HSA) fusion proteins are useful for diagnosing, treating and preventing various disorders in mammals, preferably in humans. (I)-HSA fusion proteins are also useful as molecular weight markers on sodium dodecyl sulfate polyacrylamide gel electrophoresis techniques, for raising antibodies, and to test the biological activities of the Cxbl protein. (I)-HSA fusion proteins are useful for screening for molecules that bind to the Cxbl protein portion of the fusion protein. The present sequence is used in the exemplification of the present invention.

Sequence 660 AA;

Query Match 87.2%; Score 510; DB 7; Length 660;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGEENFKALVLIAPQYLOQCPEDEHVKLVNEVTEFAKTCVADESAAE 60
DB 75 DAHSEVAHFRKDLGEENFKALVLIAPQYLOQCPEDEHVKLVNEVTEFAKTCVADESAAE 134

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLVLRPEV 120
DB 135 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLVLRPEV 194

QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFFAKRYKAAFTCCQAADKAAACLLP 180
DB 195 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFFAKRYKAAFTCCQAADKAAACLLP 254

181 KLDELDEGKASSAKQRLKCSLOKQGERAFKAWAVARLSORPPKAEFAVSKLVTDLTK 240
 255 KLDELDEGKASSAKQRLKCSLOKQGERAFKAWAVARLSORPPKAEFAVSKLVTDLTK 314
 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKJKECCKEPPLLEKSHCIAEVENDEMPA 300
 315 VHTCCHGDLLECCADRADLAKYICENQDSISSKJKECCKEPPLLEKSHCIAEVENDEMPA 374
 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIYARRHPDYSVLLLRLLAKTYETTLK 360
 375 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIYARRHPDYSVLLLRLLAKTYETTLK 434
 361 CAADPHCYAKVDFEPLVIEBPQNLKQNCBFLPQGLGEYKFNALLVRYTKVPQVST 420
 435 CAADPHCYAKVDFEPLVIEBPQNLKQNCBFLPQGLGEYKFNALLVRYTKVPQVST 494
 421 PTLVEVSNLKGSKCKEPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 495 PTLVEVSNLKGSKCKEPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 554
 481 LVNRRPCFSALEVDYVPKEFNAETTFH 510
 555 LVNRRPCFSALEVDYVPKEFNAETTFH 584

RESULT 71
 ADD06554
 ID ADD06554 standard; protein; 660 AA.
 AC ADD06554;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE Human Ckbl-HAS fusion protein construct secreted protein SEQ ID NO:90.
 KW human; chemokine betal; Ckbl; anti-HIV; neuroprotective; antithyroid;
 KW antiarthritic; antineumatic; immunosuppressive; nootropic;
 KW nephrotoxic; tuberculosic; virucide; antiallergic; osteopathic;
 KW antimicrobial; infection; HIV; immune disorder; haematopoietic disorder;
 KW autoimmune disorder; multiple sclerosis; Grave's disease; arthritis;
 KW rheumatoid arthritis; transplant rejection; neurodegenerative disorder;
 KW Alzheimer's disease; inflammatory disease; asthma; allergic disorder;
 KW inflammatory bowel disease; osteoarthritis; colitis;
 KW inflammatory kidney disease; glomerulonephritis; infectious disease;
 KW tuberculosic; hepatitis infection; herpes viral infection;
 KW viral infection; proliferative disorder; atherosclerosis;
 KW human serum albumin; HSA.
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO200297038-A2.
 XX
 PD 05-DEC-2002.
 XX
 PF 24-MAY-2002; 2002WO-US016525.
 XX
 PR 25-MAY-2001; 2001US-0293212P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Bell A, Ruben SM;
 XX
 DR WPI; 2003-140456/13.
 DR N-PSDB; ADD06552.
 XX
 PT Novel human chemokine betal protein comprising deletion in amino acids
 PT from amino and/or carboxy terminus, and is a fusion protein further
 PT comprising human serum albumin, is useful for treating multiple
 PT sclerosis, asthma.

PS Example 1; SEQ ID NO 90; 423pp; English.
 XX The present invention describes a human chemokine betal (Ckbl) protein
 CC (I) comprising a deletion in amino acid residues from the amino terminus
 CC and/or carboxy terminus of the 93 residue amino acid sequence (SI, see
 CC ADD06466). (I) has anti-HIV, neuroprotective, antithyroid, antiarthritic,
 CC antineumatic, immunosuppressive, nootropic, antiinflammatory,
 CC antiasthmatic, allergic, osteopathic, nephrotoxic, tuberculosic,
 CC virucide, antiatherosclerotic and antimicrobial activities. (I) is useful
 CC for preventing infection, preferably viral (human immunodeficiency virus
 CC (HIV)) infection, in a cell, by contacting the cell with (I). (I) is also
 CC useful for treating a disease, such as HIV infection or immune disorders,
 CC haematopoietic disorders, autoimmune disorders, multiple sclerosis,
 CC Grave's disease, arthritis, rheumatoid arthritis, transplant rejection,
 CC neurodegenerative disorders, Alzheimer's disease, inflammatory disease,
 CC asthma, allergic disorders, inflammatory bowel disease, osteoarthritis,
 CC colitis, inflammatory kidney diseases, glomerulonephritis, infectious
 CC disease, tuberculosic, hepatitis infections, herpes viral infection,
 CC viral infection, proliferative disorders or atherosclerosis, in an
 CC individual. (I) inhibits or abolishes the ability of HIV to bind to,
 CC enter into/fuse with (infect), and/or replicate in CCR5 expressing cells.
 CC (I) also acts as CCR5 agonists or antagonists, stimulate chemotaxis of
 CC CCR5-expressing cells, inhibit CCR5 ligand binding to a CCR5 molecule, or
 CC upregulate or downregulate CCR5 expression. (I) is useful as an
 CC immunological probe for the differential identification of the tissues or
 CC cell-types. (I)-human serum albumin (HSA) fusion proteins are useful for
 CC diagnosing, treating and preventing various disorders in mammals,
 CC preferably in humans. (I)-HSA fusion proteins are also useful as
 CC molecular weight markers on sodium dodecyl sulfate polyacrylamide gel
 CC electrophoresis techniques, for raising antibodies, and to test the
 CC biological activities of the Ckbl protein. (I)-HSA fusion proteins are
 CC useful for screening for molecules that bind to the Ckbl protein portion
 CC of the fusion protein. The present sequence is used in the
 CC exemplification of the present invention.
 XX
 SQ Sequence 660 AA;
 Query Match 87.2%; Score 510; DB 7; Length 660;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHSEVAHRRKIDGSENFKALVIAFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60
 DB 75 DAHSEVAHRRKIDGSENFKALVIAFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 134
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
 DB 135 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 194
 QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFPKRYKAAFTCCQAAADKACLLP 180
 DB 195 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFPKRYKAAFTCCQAAADKACLLP 254
 QY 181 KLDELDEGKASSAKQRLKCSLOKQGERAFKAWAVARLSORPPKAEFAVSKLVTDLTK 240
 DB 255 KLDELDEGKASSAKQRLKCSLOKQGERAFKAWAVARLSORPPKAEFAVSKLVTDLTK 314
 QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKJKECCKEPPLLEKSHCIAEVENDEMPA 300
 DB 315 VHTCCHGDLLECCADRADLAKYICENQDSISSKJKECCKEPPLLEKSHCIAEVENDEMPA 374
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIYARRHPDYSVLLLRLLAKTYETTLK 360
 DB 375 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIYARRHPDYSVLLLRLLAKTYETTLK 434
 QY 361 CAADPHCYAKVDFEPLVIEBPQNLKQNCBFLPQGLGEYKFNALLVRYTKVPQVST 420
 DB 435 CAADPHCYAKVDFEPLVIEBPQNLKQNCBFLPQGLGEYKFNALLVRYTKVPQVST 494
 QY 421 PTLVEVSNLKGSKCKEPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 DB 495 PTLVEVSNLKGSKCKEPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 554

QY 481 LVNRRPCFSALEVDVTVPKFNAETTFPH 510
DB 555 LVNRRPCFSALEVDVTVPKFNAETTFPH 584

RESULT 72
ADD06568
ID ADD06568 standard; protein; 676 AA.
XX
AC ADD06568;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human Cxbl-HAS fusion protein construct protein SEQ ID NO:104.

XX human; chemokine betal; Cxbl; anti-HIV; neuroprotective; antithyroid;
XX antiarthritic; antirheumatic; immunosuppressive; neutropenic;
XX antiinflammatory; antisthmatic; antiallergic; osteopathic;
XX nephrotrophic; tuberculosic; virucide; antiatherosclerotic;
XX antimicrobial; infection; HIV; immune disorder; haematopoietic disorder;
XX autoimmune disorder; multiple sclerosis; Grave's disease; arthritis;
XX rheumatoid arthritis; transplant rejection; neurodegenerative disorder;
XX Alzheimer's disease; inflammatory disease; asthma; allergic disorder;
XX inflammatory bowel disease; osteoarthritis; colitis;
XX inflammatory kidney disease; glomerulonephritis; infectious disease;
XX tuberculosis; hepatitis infection; herpes viral infection;
XX viral infection; proliferative disorder; atherosclerosis;
XX human serum albumin; HSA.

OS Synthetic.
OS Homo sapiens.

XX WO200207038-A2.

XX 05-DEC-2002.

XX 24-MAY-2002; 2002WO-US016525.

XX 25-MAY-2001; 2001US-0293212P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Bell A, Ruben SM;

XX WPI: 2003-140456/13.

XX N-PSDB; ADD06567.

XX Novel human chemokine betal protein comprising deletion in amino acids
PT from amino and/or carboxy terminus, and is a fusion protein further
PT comprising human serum albumin, is useful for treating multiple
PT sclerosis, asthma.

PS Example 1; SEQ ID NO 104; 423pp; English.

XX The present invention describes a human chemokine betal (Cxbl) protein
CC (I) comprising a deletion in amino acid residues from the amino terminus
CC and/or carboxy terminus of the 93 residue amino acid sequence (SI, see
CC ADD06466). (I) has anti-HIV, neuroprotective, antithyroid, antiaarthritic,
CC antirheumatic, immunosuppressive, neutropenic, antiatherosclerotic,
CC antisthmatic, antiallergic, osteopathic, nephrotrophic, tuberculosic,
CC virucide, antiatherosclerotic and antimicrobial activities (I) is useful
CC for preventing infection, preferably viral (human immunodeficiency virus
CC (HIV)) infection, in a cell, by contacting the cell with (I). (I) is also
CC useful for treating a disease, such as HIV infection or immune disorders,
CC haematopoietic disorders, autoimmune disorders, multiple sclerosis,
CC Grave's disease, arthritis, rheumatoid arthritis, transplant rejection,
CC neurodegenerative disorders, inflammatory bowel disease, inflammatory disease,
CC asthma, allergic disorders, inflammatory kidney diseases, glomerulonephritis, infectious
CC colitis, inflammatory kidney diseases, hepatitis infections, herpes viral infection,
CC disease, tuberculosis, proliferative disorders or atherosclerosis, in an
CC viral infection, inhibits or abolishes the ability of HIV to bind to,
CC individual. (I) inhibits or abolishes the ability of HIV to bind to,
CC enter into/fuse with (infect), and/or replicate in CCR5 expressing cells.

CC (I) also acts a CCR5 agonists or antagonists, stimulate chemotaxis of
CC CCR5-expressing cells, inhibit CCR5 ligand binding to a CCR5 molecule, or
CC upregulate or downregulate CCR5 expression. (I) is useful as an
CC immunological probe for the differential identification of the tissues or
CC cell-types. (I)-human serum albumin (HSA) fusion proteins are useful for
CC diagnosing, treating and preventing various disorders in mammals,
CC preferably in humans. (I)-HSA fusion proteins are also useful as
CC molecular weight markers on sodium dodecyl sulfate polyacrylamide gel
CC electrophoresis techniques, for raising antibodies, and to test the
CC biological activities of the Cxbl protein. (I)-HSA fusion proteins are
CC useful for screening for molecules that bind to the Cxbl protein portion
CC of the fusion protein. The present sequence is used in the
CC exemplification of the present invention.

XX Sequence 676 AA;

QY Query Match 87.2%; Score 510; DB 7; Length 676;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFKDLGRENFKALVLIAPAYLOCCPEDHVKLVNVEVFAKTCVADESAE 60
DB 91 DAHSEVAHFKDLGRENFKALVLIAPAYLOCCPEDHVKLVNVEVFAKTCVADESAE 150

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQKPERNECFLOHKDDNPNLPRLVRPEV 120
DB 151 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQKPERNECFLOHKDDNPNLPRLVRPEV 210

QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 211 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 270

QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
DB 271 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK 330

QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDENFA 300
DB 331 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDENFA 390

QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIYARRHPDYSVLLRLAKYVETTLK 360
DB 391 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIYARRHPDYSVLLRLAKYVETTLK 450

QY 361 CAADPHECYAKVDFEKLVEEPQNLKQNCLEFQGLGEYKFNQALLVRYTKVQVST 420
DB 451 CAADPHECYAKVDFEKLVEEPQNLKQNCLEFQGLGEYKFNQALLVRYTKVQVST 510

QY 421 PTLVESENGLKGVSKCKEPEAKWPCAEADYLSVLNQLCVLHEKTPVSDRVTKCTES 480
DB 511 PTLVESENGLKGVSKCKEPEAKWPCAEADYLSVLNQLCVLHEKTPVSDRVTKCTES 570

QY 481 LVNRRPCFSALEVDVTVPKFNAETTFPH 510

DB 571 LVNRRPCFSALEVDVTVPKFNAETTFPH 600

RESULT 73
ADD06559
ID ADD06559 standard; protein; 676 AA.

XX ADD06559;
AC
AC
AC
DT 01-JAN-2004 (first entry)
XX
DE Human Cxbl-HAS fusion protein construct protein SEQ ID NO:95.

XX human; chemokine betal; Cxbl; anti-HIV; neuroprotective; antithyroid;
XX antiaarthritic; antirheumatic; immunosuppressive; neutropenic;
XX antiinflammatory; antisthmatic; antiallergic; osteopathic;
XX nephrotrophic; tuberculosic; virucide; antiatherosclerotic;
XX antimicrobial; infection; HIV; immune disorder; haematopoietic disorder;
XX autoimmune disorder; multiple sclerosis; Grave's disease; arthritis;
XX

KW rheumatoid arthritis; transplant rejection; neurodegenerative disorder;
KW Alzheimer's disease; inflammatory disease; asthma; allergic disorder;
KW inflammatory bowel disease; osteoarthritis; colitis;
KW inflammatory kidney disease; glomerulonephritis; infectious disease;
KW tuberculosis; hepatitis infection; herpes viral infection;
KW viral infection; proliferative disorder; atherosclerosis;
KW human serum albumin; HSA.
XX Synthetic.
OS Homo sapiens.
XX WO200297038-A2.
XX PD 05-DEC-2002.
XX PF 24-MAY-2002; 2002WO-US016525.
XX PR 25-MAY-2001; 2001US-0293212P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX PA Bell A, Ruben SM;
XX PI WPI: 2003-140456/13.
XX DR N-PSDB; ADD06562.
XX PT Novel human chemokine betal protein comprising deletion in amino acids
PT from amino and/or carboxy terminus, and is a fusion protein further
PT comprising human serum albumin, is useful for treating multiple
PT sclerosis, asthma.
XX Example 1; SEQ ID NO 95; 423pp; English.
XX The present invention describes a human chemokine betal (Ckb1) protein.
CC (I) comprising a deletion in amino acid residues from the amino terminus
CC and/or carboxy terminus of the 93 residue amino acid sequence (SI, see
CC ADD06466). (II) has anti-HIV, neuroprotective, antithyroid, antiarthritic,
CC antirheumatic, immunosuppressive, nontropic, antiinflammatory,
CC antiasthmatic, antiallergic, osteopathic, nephrotrophic, tuberculostatic,
CC virucide, antiatherosclerotic and antimicrobial activities. (I) is useful
CC for preventing infection, preferably viral (human immunodeficiency virus
CC (HIV)) infection, in a cell, by contacting the cell with (I). (I) is also
CC useful for treating a disease, such as HIV infection or immune disorders,
CC haematopoietic disorders, autoimmune disorders, multiple sclerosis,
CC Grave's disease, arthritis, rheumatoid arthritis, transplant rejection,
CC neurodegenerative disorders, Alzheimer's disease, inflammatory disease,
CC asthma, allergic disorders, inflammatory bowel disease, osteoarthritis,
CC disease, tuberculosis, hepatitis infections, herpes viral infection,
CC colitis, inflammatory kidney diseases, glomerulonephritis, infectious
CC viral infection, proliferative disorders or atherosclerosis, in an
CC individual. (I) inhibits or abolishes the ability of HIV to bind to,
CC enter into/fuse with (infect), and/or replicate in CCR5 expressing cells.
CC (I) also acts a CCR5 agonists or antagonists, stimulate chemotaxis of
CC CCR5-expressing cells, inhibit CCR5 ligand binding to a CCR5 molecule, or
CC upregulate or downregulate CCR5 expression. (I) is useful as an
CC immunological probe for the differential identification of the tissues or
CC cell-types. (I)-human serum albumin (HSA) fusion proteins are useful for
CC diagnosing, treating and preventing various disorders in mammals,
CC preferably in humans. (I)-HSA fusion proteins are also useful as
CC molecular weight markers on sodium dodecyl sulfate polyacrylamide gel
CC electrophoresis techniques, for raising antibodies, and to test the
CC biological activities of the Ckb1 protein. (I)-HSA fusion proteins are
CC useful for screening for molecules that bind to the Ckb1 protein portion
CC of the fusion protein. The present sequence is used in the
CC exemplification of the present invention.
XX SQ Sequence 676 AA;

Query Match 87.2%; Score 510; DB 7; Length 676;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 DAHKSEVAHRFDLGEENFKALVLIAPAYLQCCPPEDHVKLVNEVTEFAKTCVADESAE 60

Db 91 DAHKSEVAHRFDLGEENFKALVLIAPAYLQCCPPEDHVKLVNEVTEFAKTCVADESAE 150
Qy 61 NCKSLHTLFGDKLCTVATIRETYGEMADCCAKQKPERNECFLOHKDDPNLPRLVRPEV 120
Db 151 NCKSLHTLFGDKLCTVATIRETYGEMADCCAKQKPERNECFLOHKDDPNLPRLVRPEV 210
Qy 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 211 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 270
Qy 181 KLDELDEGKASAKORLKCASLQKFGGERAFKAWAVARLSQRPFKAEFAEVSCLVDTLTK 240
Db 271 KLDELDEGKASAKORLKCASLQKFGGERAFKAWAVARLSQRPFKAEFAEVSCLVDTLTK 330
Qy 241 VHTCCCHGDLLECAADDRADLAKYICENQDSISSKLEKCEKPLEKSHGICIAEVENDMPA 300
Db 331 VHTCCCHGDLLECAADDRADLAKYICENQDSISSKLEKCEKPLEKSHGICIAEVENDMPA 390
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFELYEYARRHPDYSVLLRLAKTYETTLK 360
Db 391 DLPSLAADFVESKDVCKNYAEAKDVFGLMFELYEYARRHPDYSVLLRLAKTYETTLK 450
Qy 361 CAAADPHCEYAKVDFEPKPLVEEPQNLKONCELFQGLGEYKFNALLVRYTKVQVST 420
Db 451 CAAADPHCEYAKVDFEPKPLVEEPQNLKONCELFQGLGEYKFNALLVRYTKVQVST 510
Qy 421 PTLVEVSRLNKGKSKCKKPEAKRMPCADYLSVNLQNLVLEHTPVSDDRTKCTES 480
Db 511 PTLVEVSRLNKGKSKCKKPEAKRMPCADYLSVNLQNLVLEHTPVSDDRTKCTES 570
Qy 481 LVNRRPCFSALEVDVETVYKPEFNAETTFH 510
Db 571 LVNRRPCFSALEVDVETVYKPEFNAETTFH 600
RESULT 74
ADD06562
ID ADD06562 standard; protein; 676 AA.
XX ADD06562;
XX AC ADD06562;
XX DT 01-JAN-2004 (first entry)
DE Human Ckb1-HAS fusion protein construct protein SEQ ID NO:98.
XX human; chemokine betal; Ckb1; anti-HIV; neuroprotective; antithyroid;
KW antiarthritic; antirheumatic; immunosuppressive; nontropic;
KW antiinflammatory; antiasthmatic; antiallergic; osteopathic;
KW nephrotrophic; tuberculostatic; virucide; antiatherosclerotic;
KW antimicrobial; infection; HIV; immune disorder; haematopoietic disorder;
KW autoimmune disorder; multiple sclerosis; Grave's disease; arthritis;
KW rheumatoid arthritis; transplant rejection; neurodegenerative disorder;
KW Alzheimer's disease; inflammatory disease; asthma; allergic disorder;
KW inflammatory bowel disease; osteoarthritis; colitis;
KW inflammatory kidney disease; glomerulonephritis; infectious disease;
KW tuberculosis; hepatitis infection; herpes viral infection;
KW viral infection; proliferative disorder; atherosclerosis;
KW human serum albumin; HSA.
OS Synthetic.
OS Homo sapiens.
XX WO200297038-A2.
XX PD 05-DEC-2002.
XX PF 24-MAY-2002; 2002WO-US016525.
XX PR 25-MAY-2001; 2001US-0293212P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX PA

PI Bell A, Ruben SM;
XX WPI; 2003-140456/13.
DR N-PSDB; ADD06556.
XX Novel human chemokine betal protein comprising deletion in amino acids
PI from amino and/or carboxy terminus, and is a fusion protein further
PT comprising human serum albumin, is useful for treating multiple
PT sclerosis, asthma.
XX Example 1; SEQ ID NO 98; 423pp; English.
PS The present invention describes a human chemokine betal (Ckbl) protein
XX (I) comprising a deletion in amino acid residues from the amino terminus
CC and/or carboxy terminus of the 93 residue amino acid sequence (SI, see
CC ADD06466). (I) has anti-HIV, neuroprotective, antithyroid, antiarthritic,
CC antirheumatic, immunosuppressive, neurotropic, antinflammatory,
CC antiviral, antiallergic, osteopathic, nephrotrophic, tuberculostatic,
CC virucide, antiatherosclerotic and antimicrobial activities. (I) is useful
CC for preventing infection, preferably viral (human immunodeficiency virus
CC (HIV)) infection, in a cell, by contacting the cell with (I). (I) is also
CC useful for treating a disease, such as HIV infection or immune disorders,
CC haematopoietic disorders, autoimmune disorders, multiple sclerosis,
CC Grave's disease, arthritis, rheumatoid arthritis, transplant rejection,
CC neurodegenerative disorders, Alzheimer's disease, inflammatory disease,
CC asthma, allergic disorders, inflammatory bowel disease, osteoarthritis,
CC colitis, inflammatory kidney diseases, glomerulonephritis, infectious
CC disease, tuberculosis, hepatitis infections, herpes viral infection,
CC viral infection, proliferative disorders or atherosclerosis, in an
CC individual. (I) inhibits or abolishes the ability of HIV to bind to,
CC enter into/fuse with (infect), and/or replicate in CCR5 expressing cells.
CC (I) also acts a CCR5 agonists or antagonists, stimulate chemotaxis of
CC CCR5-expressing cells, inhibit CCR5 ligand binding to a CCR5 molecule, or
CC upregulate or downregulate CCR5 expression. (I) is useful as an
CC immunological probe for the differential identification of the tissues or
CC cell-types. (I)-human serum albumin (HSA) fusion proteins are useful for
CC diagnosing, treating and preventing various disorders in mammals,
CC preferably in humans. (I)-HSA fusion proteins are also useful as
CC molecular weight markers on sodium dodecyl sulfate polyacrylamide gel
CC electrophoresis techniques, for raising antibodies, and to test the
CC biological activities of the Ckbl protein. (I)-HSA fusion proteins are
CC useful for screening for molecules that bind to the Ckbl protein portion
CC of the fusion protein. The present sequence is used in the
CC exemplification of the present invention.
XX Sequence 676 AA;
SQ Query Match 87.2%; Score 510; DB 7; Length 676;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRFKDGEENFKALVLIATFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 60
DB 91 DAHKSEVAHRFKDGEENFKALVLIATFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 150
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLRLVRPEV 120
DB 151 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLRLVRPEV 210
QY 121 DVNCTAFHDNEETFLKKLYELIARHPHYFAPPELLFFAKRYKAFTCCQAADKAACLLP 180
DB 211 DVNCTAFHDNEETFLKKLYELIARHPHYFAPPELLFFAKRYKAFTCCQAADKAACLLP 270
QY 181 KLDELDEGKASAKORLKCASLOKFGERAKAVAVELSORFPAEVSJLVTLDTK 240
DB 271 KLDELDEGKASAKORLKCASLOKFGERAKAVAVELSORFPAEVSJLVTLDTK 330
QY 241 VHTCECHGDLLECCADRDADLAKYICENQDSISSKLKCECEKPLLEKSHCIAEVENDEMPA 300
DB 331 VHTCECHGDLLECCADRDADLAKYICENQDSISSKLKCECEKPLLEKSHCIAEVENDEMPA 390
QY 301 DLPSLAADFVSKDVCNVAEAKDVLGMLFLYVARRHPDYSVLLRLAKTYETTLKCC 360

DB 391 DLPSLAADFVSKDVCNVAEAKDVLGMLFLYVARRHPDYSVLLRLAKTYETTLKCC 450
QY 361 CAAAPHECYAKVDFEFLPVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420
DB 451 CAAAPHECYAKVDFEFLPVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 510
QY 421 PTLVEVSRNLGVGSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 511 PTLVEVSRNLGVGSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 570
QY 481 LVNRRPFCFSALEVDYVVPKEFNAETFTFH 510
DB 571 LVNRRPFCFSALEVDYVVPKEFNAETFTFH 600
RESULT 75
ADD06556
ID ADD06556 standard; protein; 684 AA.
XX ADD06556;
AC ADD06556;
XX 01-JAN-2004 (first entry)
XX Human Ckbl-HAS fusion protein construct protein SEQ ID NO:92.
XX human; chemokine betal; Ckbl; anti-HIV; neuroprotective; antithyroid;
XX antiarthritic; antirheumatic; immunosuppressive; neurotropic;
XX antinflammatory; antiasthmatic; antiallergic; osteopathic;
XX nephrotrophic; tuberculostatic; virucide; antiatherosclerotic;
XX antimicrobial; infection; HIV; immune disorder; haematopoietic disorder;
XX autoimmune disorder; multiple sclerosis; Grave's disease; arthritis;
XX rheumatoid arthritis; transplant rejection; neurodegenerative disorder;
XX Alzheimer's disease; inflammatory disease; asthma; allergic disorder;
XX inflammatory bowel disease; osteoarthritis; colitis;
XX inflammatory kidney disease; glomerulonephritis; infectious disease;
XX tuberculoisis; hepatitis infection; herpes viral infection;
XX viral infection; proliferative disorder; atherosclerosis;
XX human serum albumin; HSA.
XX Synthetic.
XX Homo sapiens.
XX WO200297038-A2.
XX 05-DEC-2002.
XX 24-MAY-2002; 2002WO-USO16525.
XX 25-MAY-2001; 2001US-0293212P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Bell A, Ruben SM;
XX WPI; 2003-140456/13.
XX N-PSDB; ADD06555.
XX Novel human chemokine betal protein comprising deletion in amino acids
XX from amino and/or carboxy terminus, and is a fusion protein further
XX comprising human serum albumin, is useful for treating multiple
XX sclerosis, asthma.
XX Example 1; SEQ ID NO 92; 423pp; English.
XX The present invention describes a human chemokine betal (Ckbl) protein
XX (I) comprising a deletion in amino acid residues from the amino terminus
XX and/or carboxy terminus of the 93 residue amino acid sequence (SI, see
XX ADD06466). (I) has anti-HIV, neuroprotective, antithyroid, antiarthritic,
XX antirheumatic, immunosuppressive, neurotropic, antinflammatory,
XX antiasthmatic, antiallergic, osteopathic, nephrotrophic, tuberculostatic,
XX virucide, antiatherosclerotic and antimicrobial activities. (I) is useful
XX for preventing infection, preferably viral (human immunodeficiency virus
XX (HIV)) infection, in a cell, by contacting the cell with (I). (I) is also

useful for treating a disease, such as HIV infection or immune disorders, haematopoietic disorders, autoimmune disorders, multiple sclerosis, Grave's disease, arthritis, rheumatoid arthritis, transplant rejection, neurodegenerative disorders, Alzheimer's disease, inflammatory disease, asthma, allergic disorders, inflammatory bowel disease, osteoarthritis, colitis, inflammatory kidney diseases, glomerulonephritis, infectious disease, tuberculosis, hepatitis infections, herpes viral infection, viral infection, proliferative disorders or atherosclerosis, in an individual. (I) inhibits or abolishes the ability of HIV to bind to, enter into/fuse with (infect), and/or replicate in CD4 expressing cells. (I) also acts a CD4 agonists or antagonists, stimulate chemotaxis of CD4-expressing cells, inhibit CD4 ligand binding to a CD4 molecule, or upregulate or downregulate CD4 expression. (I) is useful as an immunological probe for the differential identification of the tissues or cell-types. (I)-human serum albumin (HSA) fusion proteins are useful for diagnosing, treating and preventing various disorders in mammals, preferably in humans. (I)-HSA fusion proteins are also useful as molecular weight markers on sodium dodecyl sulfate polyacrylamide gel electrophoresis techniques, for raising antibodies, and to test the biological activities of the Ckb1 protein. (I)-HSA fusion proteins are useful for screening for molecules that bind to the Ckb1 protein portion of the fusion protein. The present sequence is used in the exemplification of the present invention.

Sequence 684 AA;

Query Match 87.2%; Score 510; DB 7; Length 684;

```
Query Match      87.2%; Score 510; DB 7; Length 684;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	1	DAHKEVAHRFKD	GEENFKALVLI	AFQYLQOCP	FEDHVKLVNEVTEFAKTCVADES	AE 60
DB	99	DAHKEVAHRFKD	GEENFKALVLI	AFQYLQOCP	FEDHVKLVNEVTEFAKTCVADES	AE 158
QY	61	NDCKSLHTLFGD	KLCTVATLRET	YGEADCCAQ	QEPERNECFLOHKDDNP	PLVRPREV 120
DB	159	NDCKSLHTLFGD	KLCTVATLRET	YGEADCCAQ	QEPERNECFLOHKDDNP	PLVRPREV 218
QY	121	DVMTAFHNBETFL	KKLYBIARRHPY	FYAPELLFP	AKYKAATFTECCQAADKAACLLP	188
DB	219	DVMTAFHNBETFL	KKLYBIARRHPY	FYAPELLFP	AKYKAATFTECCQAADKAACLLP	278
QY	181	KLDELDRSGK	ASSAKORUK	CASLOKFG	ERAFKAWAVARLSORPPK	AEAEVSKLVTDLTK 240
DB	279	KLDELDRSGK	ASSAKORUK	CASLOKFG	ERAFKAWAVARLSORPPK	AEAEVSKLVTDLTK 338
QY	241	VHTECHGDLLE	CADDRADLAKYI	CENQDSISSK	KECKECPKLEKSHCIAEV	NDEMPA 300
DB	339	VHTECHGDLLE	CADDRADLAKYI	CENQDSISSK	KECKECPKLEKSHCIAEV	NDEMPA 398
QY	301	DLPSLAADP	VESKOVCKNY	AEAKDVLGMFL	YEYARRHPDYSVVLRL	LRAKTYETTLKCK 360
DB	399	DLPSLAADP	VESKOVCKNY	AEAKDVLGMFL	YEYARRHPDYSVVLRL	LRAKTYETTLKCK 458
QY	361	CAAADPHECYAK	VFDEFKPLVEE	PQNTIKONCELF	EQLGKEYFQNALVRYTKY	QVPOVST 420
DB	459	CAAADPHECYAK	VFDEFKPLVEE	PQNTIKONCELF	EQLGKEYFQNALVRYTKY	QVPOVST 518
QY	421	PTLVEVSNL	GLKVGSKCCCH	PEAKMPCADYLS	VNLQVLVHEKTPVSDRV	TKCCTES 480
DB	519	PTLVEVSNL	GLKVGSKCCCH	PEAKMPCADYLS	VNLQVLVHEKTPVSDRV	TKCCTES 578
QY	481	LVNRP	PCFSALEVD	EYVPKFN	AEATFTFH 510	
DB	579	LVNRP	PCFSALEVD	EYVPKFN	AEATFTFH 608	

Search completed: April 19, 2004, 16:05:53
Job time : 68 secs

Result No.	Score	Query	Length	DB	ID	Description
1	36	100.0	585	1	US-08-153-799-14	Sequence 14, Appl
2	36	100.0	585	1	US-08-448-136A-3	Sequence 3, Appl
3	36	100.0	585	2	US-08-984-176-1	Sequence 1, Appl
4	36	100.0	585	2	US-08-702-512-2	Sequence 2, Appl
5	36	100.0	585	3	US-08-769-746-2	Sequence 2, Appl
6	36	100.0	585	4	US-10-153-064-5	Sequence 5, Appl
7	36	100.0	609	1	US-08-223-619-3	Sequence 3, Appl
8	36	100.0	609	1	US-08-433-037-4	Sequence 4, Appl
9	36	100.0	609	4	US-08-897-956A-2	Sequence 2, Appl
10	36	100.0	609	4	US-10-153-064-7	Sequence 7, Appl
11	36	100.0	609	4	US-09-576-594-977	Sequence 977, App
12	36	100.0	609	5	PCT-US95-04075-3	Sequence 3, Appl
13	36	100.0	610	2	US-08-797-689-2	Sequence 2, Appl
14	36	100.0	610	4	US-09-984-186-2	Sequence 2, Appl
15	36	100.0	651	4	US-10-153-064-133	Sequence 133, App
16	36	100.0	652	4	US-10-153-064-96	Sequence 96, Appl
17	36	100.0	652	4	US-10-153-064-99	Sequence 99, Appl
18	36	100.0	652	4	US-10-153-064-105	Sequence 105, App
19	36	100.0	652	4	US-10-153-064-132	Sequence 132, App
20	36	100.0	653	4	US-10-153-064-131	Sequence 131, App
21	36	100.0	656	4	US-10-153-064-130	Sequence 130, App
22	36	100.0	660	4	US-10-153-064-90	Sequence 90, Appl
23	36	100.0	660	4	US-10-153-064-93	Sequence 93, Appl
24	36	100.0	668	4	US-10-153-064-102	Sequence 102, App
25	36	100.0	676	4	US-10-153-064-95	Sequence 95, Appl
26	36	100.0	676	4	US-10-153-064-98	Sequence 98, Appl
27	36	100.0	676	4	US-10-153-064-104	Sequence 104, App

MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Region
LOCATION: 369..419
OTHER INFORMATION: /note= "Alternative C-termini of
OTHER INFORMATION: HSA(1-n)"
FEATURE:
NAME/KEY: Region
LOCATION: 1..585
OTHER INFORMATION: /note= "Amino acid sequence of
OTHER INFORMATION: natural HSA"
US-08-153-799-14
Query Match 100.0%; Score 36; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KADDKET 7
Db 560 KADDKET 566
RESULT 2
US-08-448-196A-3
Sequence 3, Application US/08448196A
Patent No. 5780594
GENERAL INFORMATION:
APPLICANT: CARTER, DANIEL C.
TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
TITLE OF INVENTION: RELATED PROTEINS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: NASA
STREET: MARSHALL SPACE FLIGHT CENTER
CITY: HUNTSVILLE
STATE: ALABAMA
COUNTRY: USA
ZIP: 35812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,196A
FILING DATE: 23-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: BROAD JR., ROBERT L.
REGISTRATION NUMBER: 18,757
REFERENCE/DOCKET NUMBER: XX/WFS-28402-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 205-544-0021
TELEFAX: 205-544-0258
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-448-196A-3
Query Match 100.0%; Score 36; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
Db 560 KADDKET 566
RESULT 3
US-08-984-176-1
Sequence 1, Application US/08984176
Patent No. 5948609
GENERAL INFORMATION:
APPLICANT: CARTER, DANIEL C
APPLICANT: HO, JOSEPH X
APPLICANT: RUKER, FLORIAN
TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT
TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER
FILE REFERENCE: 08/984,176
CURRENT APPLICATION NUMBER: US/08/984,176
CURRENT FILING DATE: 1997-12-03
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 585
TYPE: PRT
ORGANISM: Homo sapiens
US-08-984-176-1
Query Match 100.0%; Score 36; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KADDKET 7
Db 560 KADDKET 566
RESULT 4
US-08-702-572-2
Sequence 2, Application US/08702572
Patent No. 5965386
GENERAL INFORMATION:
APPLICANT: Kerry-Williams, Sean M
APPLICANT: Gilbert, Sarah C
TITLE OF INVENTION: Yeast Strains and Modified Albumins
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Centeon L.L.C.
STREET: 1020 First Avenue
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-1310
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: MicroSoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,572
FILING DATE: 11-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 95/23857
FILING DATE: 1-MAR-1995
APPLICATION NUMBER: GB 9404270.2
FILING DATE: 5-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Naomi Biswas
REGISTRATION NUMBER: 38,384
REFERENCE/DOCKET NUMBER: CE0114 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610/878/4294
TELEFAX: 610/878/4221
INFORMATION FOR SEQ ID NO: 2:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-702-572-2

Query Match      100.0%; Score 36; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KADDKET 7
Db      560 KADDKET 566

RESULT 5
US-08-769-746-2
; Sequence 2, Application US/08769746
; Patent No. 6274305
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; APPLICANT: Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,746
; FILING DATE: 19-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MERI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-769-746-2

Query Match      100.0%; Score 36; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KADDKET 7
Db      560 KADDKET 566

RESULT 6
US-10-153-064-5
; Sequence 5, Application US/10153064
; Patent No. 6663495
; GENERAL INFORMATION:
; APPLICANT: Heil et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556

; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-832-929-18_copy_560_566.ra1

; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5

Query Match      100.0%; Score 36; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KADDKET 7
Db      560 KADDKET 566

RESULT 7
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri
; APPLICANT: Lyons, David
; APPLICANT: Wurfel, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RRC
; STREET: 1840 Behavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,619
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-222-619-3

Query Match      100.0%; Score 36; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KADDKET 7
Db      584 KADDKET 590

RESULT 8
US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828
; GENERAL INFORMATION:
; APPLICANT: Sreekrishna, Kotikanyadan
```

; APPLICANT: Barr, Kathryn A.
 ; APPLICANT: Brierley, Russell A.
 ; APPLICANT: Thill, Gregory F.
 ; APPLICANT: Tschoop, Jueis F.
 ; TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN
 ; TITLE OF INVENTION: PICHIA PASTORIS
 ; NUMBER OF SEQUENCES: 19
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Scully, Scott, Murphy & Presser
 ; STREET: 400 Garden City Plaza
 ; CITY: Garden City
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 11530-0299
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/433,037
 ; FILING DATE: 03-MAY-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Digiglio, Frank S.
 ; REGISTRATION NUMBER: 31,346
 ; REFERENCE/DOCKET NUMBER: 91082
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (516) 742-4343
 ; TELEFAX: (516) 742-4366
 ; TELEX: 230 901 SANS UR
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 609 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-433-037-4

Query Match 100.0%; Score 36; DB 1; Length 609;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDDKET 7
 DB 584 KADDDKET 590

RESULT 9
 US-08-897-956A-2
 ; Sequence 2, Application US/08897956A
 ; Patent No. 6423512
 ; GENERAL INFORMATION:
 ; APPLICANT: Mary Ellen Digan
 ; APPLICANT: Philip Lake
 ; APPLICANT: Hermann Gram
 ; TITLE OF INVENTION: Fusion Polypeptides
 ; FILE REFERENCE: 600-7244/CPA
 ; CURRENT APPLICATION NUMBER: US/08/897,956A
 ; CURRENT FILING DATE: 1997-07-21
 ; PRIOR APPLICATION NUMBER: 60/022,689
 ; PRIOR FILING DATE: 1996-07-26
 ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 609
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 ; US-08-897-956A-2

Query Match 100.0%; Score 36; DB 4; Length 609;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDDKET 7
 DB 584 KADDDKET 590

RESULT 10
 US-10-153-064-7
 ; Sequence 7, Application US/10153064
 ; Patent No. 6663485
 ; GENERAL INFORMATION:
 ; APPLICANT: Bell et al.
 ; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
 ; FILE REFERENCE: PFS56
 ; CURRENT APPLICATION NUMBER: US/10/153,064
 ; CURRENT FILING DATE: 2002-05-24
 ; PRIOR APPLICATION NUMBER: 60/293,212
 ; PRIOR FILING DATE: 2001-05-25
 ; NUMBER OF SEQ ID NOS: 137
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 7
 ; LENGTH: 609
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 ; US-10-153-064-7

Query Match 100.0%; Score 36; DB 4; Length 609;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDDKET 7
 DB 584 KADDDKET 590

RESULT 11
 US-09-976-594-977
 ; Sequence 977, Application US/09976594
 ; Patent No. 6673549
 ; GENERAL INFORMATION:
 ; APPLICANT: Furness, Michael
 ; APPLICANT: Buchbinder, Jenny
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
 ; FILE REFERENCE: PA-0041 US
 ; CURRENT APPLICATION NUMBER: US/09/976,594
 ; CURRENT FILING DATE: 2001-10-12
 ; PRIOR APPLICATION NUMBER: 60/240,409
 ; PRIOR FILING DATE: 2000-10-12
 ; NUMBER OF SEQ ID NOS: 1143
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 977
 ; LENGTH: 609
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. 6673549 088957CD1
 ; US-09-976-594-977

Query Match 100.0%; Score 36; DB 4; Length 609;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDDKET 7
 DB 584 KADDDKET 590

RESULT 12
 PCT-US95-04075-3
 ; Sequence 3, Application PC/TUS9504075
 ; GENERAL INFORMATION:
 ; APPLICANT: AMGEN INC.

;; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
;; TITLE OF INVENTION: Protein
;; NUMBER OF SEQUENCES: 33
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Amgen Center, Patent Operations/RRC
;; STREET: 1840 DeHavilland Drive
;; CITY: Thousand Oaks
;; STATE: California
;; COUNTRY: U.S.
;; ZIP: 91320-1789
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US95/04075
;; FILING DATE:
;; CLASSIFICATION:
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 609 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: unknown
;; TOPOLOGY: unknown
;; MOLECULE TYPE: protein
PCT-US95-04075-3

Query Match 100.0%; Score 36; DB 5; Length 609;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
|||||
DB 584 KADDKET 590

RESULT 13
US-08-797-689-2
; Sequence 2, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guittou, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
PCT-US95-04075-3

;; APPLICATION NUMBER: PCT/FR93/00085
;; FILING DATE: 28-JAN-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Smith Ph.D., Julie K.
;; REGISTRATION NUMBER: P-38,619
;; REFERENCE/DOCKET NUMBER: ST92006-US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (610) 454-3839
;; TELEFAX: (610) 454-3808
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 610 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-797-689-2

Query Match 100.0%; Score 36; DB 2; Length 610;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
|||||
DB 584 KADDKET 590

RESULT 14
US-09-984-186-2
; Sequence 2, Application US/09984186
; Patent No. 6686179
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guittou, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,186
; FILING DATE: 29-OCT-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-797-689-2

; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-984-186-2

Query Match 100.0%; Score 36; DB 4; Length 610;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
Db 584 KADDKET 590

RESULT 15
US-10-153-064-133
; Sequence 133, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 133
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-133

Query Match 100.0%; Score 36; DB 4; Length 651;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
Db 626 KADDKET 632

Search completed: April 19, 2004, 12:05:21
Job time : 1.07295 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 16:03:40 ; Search time 20 Seconds
(without alignments)
2813.602 Million cell updates/sec

Title: US-09-832-929-18

Perfect score: 585

Sequence: 1 DAHKSEVARRFKDLGNEFK.....TCFAEKGKLVAAASQAALGL 585

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

PIR 78:.*
1: PIR1:.*
2: PIR2:.*
3: PIR3:.*
4: PIR4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	585	100.0	609	1	serum albumin prec
2	75	12.8	600	2	serum albumin prec
3	29	5.0	608	2	serum albumin prec
4	27	4.6	607	1	serum albumin prec
5	26	4.4	605	1	serum albumin prec
6	26	4.4	608	1	serum albumin prec
7	25	4.3	609	2	serum albumin prec
8	22	3.8	607	1	serum albumin prec
9	22	3.8	607	1	serum albumin prec
10	20	3.4	265	2	albumin - dog (fra
11	18	3.1	453	2	serum albumin - no
12	8	1.4	453	2	serum albumin - do
13	8	1.4	273	2	serum albumin - do
14	8	1.4	289	2	serum albumin - do
15	8	1.4	538	1	serum albumin - do
16	8	1.4	568	1	serum albumin - do
17	8	1.4	575	1	serum albumin - do
18	8	1.4	575	1	serum albumin - do
19	8	1.4	575	1	serum albumin - do
20	8	1.4	575	1	serum albumin - do
21	8	1.4	609	1	serum albumin - do
22	8	1.4	609	1	serum albumin - do
23	8	1.4	609	1	serum albumin - do
24	8	1.4	615	1	serum albumin - do
25	8	1.4	760	2	serum albumin - do
26	8	1.4	1090	2	serum albumin - do
27	8	1.4	1169	2	serum albumin - do
28	7	1.2	77	2	serum albumin - do
29	7	1.2	122	2	serum albumin - do

30	7	1.2	123	2	S41361	genome polypeptide
31	7	1.2	124	2	S41360	genome polypeptide
32	7	1.2	125	2	S41352	genome polypeptide
33	7	1.2	126	2	AC2901	hypothetical prote
34	7	1.2	126	2	E97676	succinate dehydrog
35	7	1.2	126	2	S70820	H+-transporting tw
36	7	1.2	126	2	B83835	hypothetical prote
37	7	1.2	126	2	F82675	transcription regu
38	7	1.2	126	2	T45261	hypothetical prote
39	7	1.2	126	2	B82521	hypothetical prote
40	7	1.2	126	2	PS0388	genome polypeptide
41	7	1.2	126	2	S33740	polyprotein - hepa
42	7	1.2	126	2	C69266	NAD(P)H-flavin oxi
43	7	1.2	126	2	B83910	hypothetical prote
44	7	1.2	126	2	A70763	hypothetical prote
45	7	1.2	126	2	C71691	ribosomal protein
46	7	1.2	126	2	D97758	30S ribosomal prot
47	7	1.2	126	2	T46434	hypothetical prote
48	7	1.2	126	2	F70708	hypothetical prote
49	7	1.2	126	2	E69732	PBSX prophage ORF
50	7	1.2	126	2	B69020	conserved hypothet
51	7	1.2	126	2	S43591	M04D8.4 protein -
52	7	1.2	126	2	S77870	dnak-type molecula
53	7	1.2	126	2	I29034	hypothetical prote
54	7	1.2	126	2	S64849	hypothetical prote
55	7	1.2	126	2	PQ0393	genome polypeptide
56	7	1.2	126	2	AG1037	probable capsid pr
57	7	1.2	126	2	S30124	hypothetical prote
58	7	1.2	126	2	T19411	hypothetical prote
59	7	1.2	126	2	B81404	probable transcrip
60	7	1.2	126	2	T05510	probable glutathio
61	7	1.2	126	2	E71042	hypothetical prote
62	7	1.2	126	2	AC0258	probable phage pro
63	7	1.2	126	2	AE3585	ferric anguibactin
64	7	1.2	126	2	UN0265	genome polypeptide
65	7	1.2	126	2	A54932	zeta-crystallin /
66	7	1.2	126	2	D81787	probable branched-
67	7	1.2	126	2	I55593	connexin 37 - huma
68	7	1.2	126	2	T11785	adenosine deaminas
69	7	1.2	126	2	D81352	dihydroorotate oxi
70	7	1.2	126	2	T30754	probable 3beta-hyd
71	7	1.2	126	2	A48931	transmembrane glyc
72	7	1.2	126	2	AE3175	hypothetical prote
73	7	1.2	126	2	C72590	probable hexosyltr
74	7	1.2	126	2	S43159	outer membrane por
75	7	1.2	126	2	AE0615	outer membrane pro
76	7	1.2	126	2	E75012	hypothetical prote
77	7	1.2	126	2	S24771	genome polypeptide
78	7	1.2	126	2	T05124	dihydroorotase (EC
79	7	1.2	126	2	T24148	hypothetical prote
80	7	1.2	126	2	AD1559	glycine betaine AB
81	7	1.2	126	2	AF1201	glycine betaine AB
82	7	1.2	126	2	S74528	hypothetical prote
83	7	1.2	126	2	REMS5	renin (EC 3.4.23.1
84	7	1.2	126	2	AP2497	transposase all715
85	7	1.2	126	2	FC3060	genome polypeptide
86	7	1.2	126	2	FC3061	genome polypeptide
87	7	1.2	126	2	F87709	DNA/pantothenate m
88	7	1.2	126	2	A70214	hypothetical prote
89	7	1.2	126	2	A42843	heat shock protein
90	7	1.2	126	2	A97300	gamma-glutamyl pho
91	7	1.2	126	2	D42725	nitrite hydratase
92	7	1.2	126	2	D85911	probable exported
93	7	1.2	126	2	E87508	outer membrane pro
94	7	1.2	126	2	S12707	genome polypeptide
95	7	1.2	126	2	T18903	hypothetical prote
96	7	1.2	126	2	S41288	genome polypeptide
97	7	1.2	126	2	A82234	gonadolibelin III-
98	7	1.2	126	2	T46852	anthranilate synth
99	7	1.2	126	2	PC1284	genome polypeptide
100	7	1.2	126	2	A44150	structural protein

ALIGNMENTS

RESULT 1

ABHUS
serum albumin precursor [validated] - human
N/Alternate names: preproalbumin
N/Contains: kinetensin
C/Species: Homo sapiens (man)
C/Date: 29-Jul-1991 #sequence revision 31-Jan-1997 #text change 17-Mar-2000
C/Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S36
R/Lawn, R.W.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebur
Nucleic Acids Res. 9, 6103-6114, 1981
A/Title: The sequence of human serum albumin cDNA and its expression in *Escherichia coli*
A/Reference number: A93743; MUID:82081882; PMID:6171778
A/Accession: A93743
A/Molecule type: mRNA
A/Residues: 1-419, 'K', 122-609 <LAW>
A/Cross-references: EMBL:V00495; GB:J00078; GB:L00132; GB:L00133; NID:G28591; PIDN:CAA23
R/Dugaiczky, A.; Law, S.W.; Denison, O.E.
Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982
A/Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.
A/Reference number: A93936; MUID:82105994; PMID:6275391
A/Accession: A93936
A/Molecule type: mRNA
A/Residues: 1-120, 'G', 122-609 <DUG>
A/Cross-references: EMBL:V00494; NID:G28589; PIDN:CAA23753.1; PID:G28590
R/Urano, Y.; Watanabe, K.; Sakai, M.; Tanooki, T.
J. Biol. Chem. 261, 3244-3251, 1986
A/Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and
A/Reference number: I39427; MUID:86140099; PMID:2419329
A/Accession: I39427
A/Status: translation not shown
A/Molecule type: DNA
A/Residues: 1-26 <URA>
A/Cross-references: GB:ML3075; NID:G178330; PIDN:AAAS1688.1; PID:G553173
R/Watkins, J.; Madson, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994
A/Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family.
A/Reference number: I59286; MUID:94181575; PMID:8134387
A/Accession: I59286
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 282-290, 'KSRPDLQ' <WAT>
A/Cross-references: GB:S69192; NID:G546032; PIDN:AAB30282.1; PID:G546033
A/Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia
R/Madison, J.; Galliano, M.; Watkins, J.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam, F.
Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994
A/Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl-
A/Reference number: I59313; MUID:94294404; PMID:8022807
A/Accession: I59313
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 589-590, 'ALPRRVKMLLOVKLP' <MAD>
A/Cross-references: GB:S70799; NID:G547231; PIDN:AAB31177.1; PID:G547232
A/Note: this frame-shift variant is designated albumin Bazzano; four additional variants
R/Menaya, J.; Parrilla, R.; Ayuso, M.S.
submitted to the EMBL Data Library, March 1995
A/Reference number: G08292
A/Accession: G01747
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-120, 'G', 122-455 <MEN>
A/Cross-references: EMBL:U02961; NID:G763428; PIDN:AAA64922.1; PID:G763431
R/Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.
Biochem. J. 308, 321-325, 1995
A/Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast *Kex2*
A/Reference number: S55314; MUID:95275251; PMID:7755581
A/Accession: S55314
A/Molecule type: protein
A/Residues: 19-27 <LED>
R/Meloun, B.; Moravsek, L.; Kostka, V.
FEBS Lett. 58, 134-137, 1975
A/Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported
A/Title: Complete amino acid sequence of human serum albumin.
A/Reference number: A91420; MUID:76187907; PMID:1225573
A/Accession: A91420
A/Molecule type: protein
A/Residues: 25-117, 'EQ', 120-154, 'Q', 156-193, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395
R/Roehr, U.; Spittler, G.; Tripiet, D.
Justus Liebig's Ann. Chem. 9, 881-884, 1988
A/Title: Isolation and structure elucidation of middle-molecular weight peptides from u
A/Reference number: S06422
A/Note: this paper is in German, with an English abstract
A/Accession: S06422
A/Molecule type: protein
A/Residues: 25-48 <ROE>
R/Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.
Arch. Biochem. Biophys. 305, 595-599, 1993
A/Title: Mass spectrometric identification of modifications to human serum albumin trea
A/Reference number: S36882; MUID:93384321; PMID:8373198
A/Accession: S36882
A/Molecule type: protein
A/Residues: 45-67; 141-160; 311-337; 469-490; 570-581 <FIN>
R/Kausler, E.; Spittler, G.
Biol. Chem. Hoppe-Seyler 372, 849-855, 1991
A/Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmc
A/Reference number: S17599; MUID:92126241; PMID:1773598
A/Accession: S17599
A/Molecule type: protein
A/Residues: 25-54; 354-357; 431-447 <KAU>
A/Note: 49-Leu was also found
R/Caraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A/Title: Structures of histamine-releasing peptides formed by the action of acid protea
A/Reference number: A45800; MUID:89341406; PMID:2474609
A/Accession: A45800
A/Molecule type: protein
A/Residues: 166-173, 'L' <MOG>
R/McGard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; W
Biochem. Biophys. Res. Commun. 136, 983-988, 1986
A/Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-tr
A/Reference number: A03239; MUID:86242180; PMID:3067352
A/Accession: A03239
A/Molecule type: protein
A/Residues: 166-173, 'L' <MOG>
R/Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, J.
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990
A/Title: Mutations in genetic variants of human serum albumin found in Italy.
A/Reference number: A38255; MUID:91062352; PMID:2247440
A/Accession: A38255
A/Molecule type: protein
A/Residues: 76-111 <GAL1>
A/Accession: B38255
A/Molecule type: protein
A/Residues: 82-105, 'K', 107-110 <GAL2>
A/Note: this variant is designated albumin Vibo Valentia
A/Accession: A38255
A/Molecule type: protein
A/Residues: 76-83, 'K', 85-106 <GAL3>
A/Note: this variant is designated albumin Torino
R/Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.
Eur. J. Biochem. 214, 437-444, 1993
A/Title: The structural characterization and bilirubin-binding properties of albumin He
A/Reference number: S33298; MUID:93292504; PMID:8513793
A/Accession: S33298
A/Molecule type: protein
A/Residues: 255-263, 'E', 265-281 <MIN1>
A/Note: this variant is designated albumin Herborn
R/Minchiotti, L.; Galliano, M.; Stoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta
Biochem. Biophys. Acta 1119, 232-238, 1992
A/Title: Two albumins with identical electrophoretic mobility are produced by diffe
A/Reference number: S21078; MUID:92190239; PMID:1347703
A/Accession: S21078
A/Molecule type: protein
A/Residues: 354-356, 'K', 358-378 <MIN2>
A/Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported

C:Species: Felis silvestris catus (domestic cat)
C>Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Aug-1999
C:Accession: J04660; S57632
R:Weinstock, J.; Grigioni, F.; Hentges, F.
Gene 169, 295-296, 1996
A:Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.
A:Reference number: J04660; MUID:96194824; PMID:8647469
A:Accession: J04660
A:Molecule type: mRNA
A:Residues: 1-608 <HT2>
A:Cross-references: EMBL:X84842; NID:9886484; PIDN:CAA59279.1; PID:9886485
A:Experimental source: liver
A:Comment: This protein is the major protein component in plasma. It functions as a mult
e in has 35 conserved cysteine residues.
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: liver; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PR>
F:25-608/Product: serum albumin #status predicted <WAT>
F:29-202/Domain: serum albumin repeat homology <SA1>
F:231-394/Domain: serum albumin repeat homology <SA2>
F:412-592/Domain: serum albumin repeat homology <SA3>
F:261/Binding site: bilirubin (Lys) #status predicted

Query Match 5.0%; Score 29; DB 2; Length 608;
Best Local Similarity 100.0%; Pred. No. 1.8e-20;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 244 ECCHGDLLECCADRADLAKYICENQDSIS 272
Db 269 ECCHGDLLECCADRADLAKYICENQDSIS 296

RESULT 4
ABHOS
serum albumin precursor - horse
C:Species: Equus caballus (domestic horse)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: S34053
R:Ho, J.X.; Holowachuk, E.W.; Norton, E.J.; Twigg, P.D.; Carter, D.C.
Eur. J. Biochem. 215, 205-212, 1993
A:Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm
A:Reference number: S34053; MUID:93345495; PMID:8344282
A:Accession: S34053
A:Molecule type: mRNA
A:Residues: 1-607 <HOA>
A:Cross-references: GB:X74045; NID:939671; PIDN:CAA52194.1; PID:939672
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the membra
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-607/Product: serum albumin #status predicted <WAT>
F:29-201/Domain: serum albumin repeat homology <SA1>
F:412-591/Domain: serum albumin repeat homology <SA2>
F:27/Binding site: copper (His) #status predicted <SA3>
F:77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4
F:263/Binding site: bilirubin (Lys) #status predicted

Query Match 4.6%; Score 27; DB 1; Length 607;
Best Local Similarity 100.0%; Pred. No. 1.8e-18;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 55 ADESAENCXSLHTLFGDKLCTVATLR 81
Db 79 ADESAENCXSLHTLFGDKLCTVATLR 105

RESULT 5
ABPGS
serum albumin precursor - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)

C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: S01382; A61006
R:Weinstock, J.; Baldwin, G.S.
Nucleic Acids Res. 16, 9045, 1988
A:Title: Nucleotide sequence of porcine liver albumin.
A:Reference number: S01382; MUID:89016582; PMID:3174440
A:Accession: S01382
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-605 <WEI>
A:Cross-references: EMBL:X12422; NID:91875; PIDN:CAA30970.1; PID:9833798
R:Lineback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.
J. Bone Miner. Res. 4, 235-241, 1989
A:Title: Serum albumin and its acid hydrolysis peptides dominate preparations of minera
A:Reference number: A61006; MUID:89269769; PMID:2728927
A:Accession: A61006
A:Molecule type: protein
A:Residues: 23-53, 'X', 53-54; 'XXGX', 146, 'E', 148, 'E', 150-151, 'XVX', 155 <LIM>
A:Experimental source: dental enamel
A:Note: albumin and other serum proteins are also found in bone
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the membr
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-16/Domain: signal sequence (fragment) #status predicted <SIG>
F:17-22/Domain: propeptide #status predicted <PRO>
F:23-605/Product: serum albumin #status predicted <WAT>
F:27-193/Domain: serum albumin repeat homology <SA1>
F:218-391/Domain: serum albumin repeat homology <SA2>
F:410-589/Domain: serum albumin repeat homology <SA3>
F:75-84,97-113,112-123,145-190,189-198,221-267,266-274,286-300,299-310,337-362,381-390,
F:261/Binding site: bilirubin (Lys) #status predicted

Query Match 4.4%; Score 26; DB 1; Length 605;
Best Local Similarity 100.0%; Pred. No. 1.9e-17;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 244 ECCHGDLLECCADRADLAKYICENQD 269
Db 265 ECCHGDLLECCADRADLAKYICENQD 290

RESULT 6
ABRTS
serum albumin precursor - rat
N:Alternate names: preproalbumin
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-May-1979 #sequence_revision 31-May-1979 #text_change 22-Jun-1999
C:Accession: A93872; A92211; A91946; A91940; C45800; I57621; A03233
R:Sargent, T.D.; Yang, M.; Bonner, J.
Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981
A:Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.
A:Reference number: A93872; MUID:81223722; PMID:7017712
A:Accession: A93872
A:Molecule type: mRNA
A:Residues: 1-608 <SAR>
A:Cross-references: GB:V01222; GB:J00698; NID:955627; PIDN:CAA24532.1; PID:955628
R:Strauss, A.W.; Bennett, C.D.; Donohue, A.M.; Rodkey, J.A.; Alberste, A.W.
J. Biol. Chem. 252, 6846-6855, 1977
A:Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analy
A:Reference number: A92211; MUID:77249657; PMID:893447
A:Note: cleavages during protein maturation
A:Accession: A92211
A:Molecule type: protein
A:Residues: 1-38 <STR>
R:Isemura, S.; Ikenaka, T.
J. Biochem. 83, 35-48, 1978
A:Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleava
A:Reference number: A91946; MUID:78109429; PMID:564345
A:Accession: A91946
A:Molecule type: protein
A:Residues: 25-222 <ISI>
R:Isemura, S.; Ikenaka, T.

J. Biochem. 79, 1183-1196, 1976

A/Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino acid sequence of the fragments

A/Reference number: A91940; MUID:76260153; PMID:956149

A/Accession: A91940

A/Molecule type: protein

A/Residues: 223-288;572-608 <1G2>

A/Note: 262-Leu was also found

R/Aoyagi, Y.; Ikenaka, T.; Tachida, F.

Cancer Res. 38, 3483-3486, 1978

A/Title: Copper(II)-binding ability of human alpha-fetoprotein.

A/Reference number: A90758; MUID:79001617; PMID:80265

A/Contents: annotation; copper binding

R/Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.

J. Immunol. 143, 1680-1684, 1989

A/Title: Structures of histamine-releasing peptides formed by the action of acid proteases

A/Reference number: A45800; MUID:89341406; PMID:2474609

A/Accession: C45800

A/Status: preliminary

A/Molecule type: protein

A/Residues: 166-173 <CAR>

R/Heard, J.

Mol. Cell. Biol. 7, 2425-2434, 1987

A/Title: Determinants of rat albumin promoter tissue specificity analyzed by an improved

A/Reference number: I57621; MUID:87286876; PMID:3475566

A/Accession: I57621

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-5 <RES>

A/Cross-references: GB:M16825; NID:g202828; PID:AAA40712.1; PID:g554412

C/Superfamily: serum albumin; serum albumin repeat homology

C/Keywords: carrier protein; duplication; metal binding; plasma

F/1-18/Domain: signal sequence; status experimental <SIG>

F/19-24/Domain: propeptide; status experimental <PRO>

F/25-608/Product: serum albumin; status experimental <MAT>

F/29-202/Domain: serum albumin repeat homology <SAI>

F/221-394/Domain: serum albumin repeat homology <SA2>

F/413-592/Domain: serum albumin repeat homology <SA3>

F/27/Binding site: copper (His) #status experimental

F/77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,4

4.4%; Score 26; DB 1; Length 608;

Query Match 100.0%; Pred. No. 1.9e-17; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 ADCCAKQEPNERCEFLQHKDDNP 113

112 ADCCAKQEPNERCEFLQHKDDNP 137

DB

RESULT 7

JC5838

albumin - Mongolian jird

C/Species: Meriones unguiculatus (Mongolian jird)

C/Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 19-May-2000

C/Accession: JC5838

R/Yoshida, K.; Seto-Ohehima, A.; Sinohara, H.

DNA Res. 4, 351-354, 1997

A/Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthesis in the

A/Reference number: JC5838; MUID:98116663; PMID:9455485

A/Accession: JC5838

A/Molecule type: mRNA

A/Residues: 1-609 <YOS>

A/Cross-references: DBJ:AB006197; NID:g2317277; PID:BAA21765.1; PID:g2317278

A/Experimental source: liver

A/Superfamily: serum albumin; serum albumin repeat homology

F/222-395/Domain: serum albumin repeat homology <SA2>

Query Match 4.3%; Score 25; DB 2; Length 609;

Best Local Similarity 100.0%; Pred. No. 1.9e-16;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 EMADCCAKQEPNERCEFLQHKDDNP 110

111 EMADCCAKQEPNERCEFLQHKDDNP 135

DB

RESULT 8

ABOS

serum albumin precursor [validated] - bovine

N/Alternate names: 67K protein; prealbumin

C/Species: Bos primigenius taurus (cattle)

C/Date: 24-Apr-1984 #sequence_revision 30-Sep-1993 #text_change 18-Aug-2000

C/Accession: A38895; A36401; A91258; B60808; S10780; A26693; A90309; A91458; A9

R/Holowachuk, E.W.; Stolteneborg, J.K.; Reed, R.G.; Peters Jr., T.

submitted to the EMBL Data Library, August 1991

A/Description: Bovine serum albumin; cDNA sequence and expression.

A/Reference number: A38885

A/Accession: A38885

A/Molecule type: mRNA

A/Residues: 1-607 <HOL>

A/Cross-references: EMBL:M73215

R/Hirayama, K.; Akashi, S.; Furuya, M.; Fukuhata, K.

Biochem. Biophys. Res. Commun. 173, 639-646, 1990

A/Title: Rapid confirmation and revision of the primary structure of bovine serum album

A/Reference number: A36401; MUID:91083649; PMID:2260975

A/Accession: A36401

A/Molecule type: protein

A/Residues: 25-41, 'H', 43-189, 'E', 191-213, 'T', 215-323, 'D', 325-393, 'TS', 396-607 <HIR>

R/MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.

Eur. J. Biochem. 98, 477-485, 1979

A/Title: Biosynthesis of bovine plasma proteins in a cell-free system.

A/Reference number: A91258; MUID:80024278; PMID:488109

A/Accession: A91258

A/Molecule type: protein

A/Residues: 1-32 <MAG>

R/Hsieh, J.C.; Lin, F.P.; Tam, M.F.

Anal. Biochem. 170, 1-8, 1988

A/Title: Electrophoretic transfer of bovine plasma proteins from an analytical isoelectrofocusing

A/Reference number: A60806; MUID:88267456; PMID:3389500

A/Accession: B60808

A/Molecule type: protein

A/Residues: 25-41 <HSI>

R/Strawich, E.; Glimcher, M.J.

Eur. J. Biochem. 191, 47-56, 1990

A/Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelins' is alb

A/Reference number: S10780; MUID:90336641; PMID:2379503

A/Accession: S10780

A/Molecule type: protein

A/Residues: 25-41, 'H', 43-57, 59-64 <STR>

R/Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.

J. Immunol. 143, 1680-1684, 1989

A/Title: Structures of histamine-releasing peptides formed by the action of acid protea

A/Reference number: A45800; MUID:89341406; PMID:2474609

A/Accession: D45800

A/Molecule type: protein

A/Residues: 163-172 <CAR>

R/Carraway, R.E.; Mitra, S.P.; Cochran, D.E.

J. Biol. Chem. 262, 5968-5973, 1987

A/Title: Structure of a biologically active neurotensin-related peptide obtained from p

A/Reference number: A26693; MUID:87194805; PMID:2437111

A/Accession: A26693

A/Molecule type: protein

A/Residues: 165-172, 'L', <CA2>

R/Reed, R.G.; Putnam, F.W.; Peters Jr., T.

Biochem. J. 191, 867-868, 1980

A/Title: Sequence of residues 400-403 of bovine serum albumin.

A/Reference number: A90309; MUID:82023364; PMID:7283978

A/Accession: A90309

A/Molecule type: protein

A/Residues: 402-433 <REE>

R/Brown, J.R.

Fed. Proc. 34, 591, 1975

A/Title: Structure of bovine serum albumin.

A/Reference number: A91458

A/Accession: A91458

A/Molecule type: protein

J. Biochem. 79, 1183-1196, 1976

A/Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino acid sequence of the fragments

A/Reference number: A91940; MUID:76260153; PMID:956149

A/Accession: A91940

A/Molecule type: protein

A/Residues: 223-288;572-608 <1G2>

A/Note: 262-Leu was also found

R/Aoyagi, Y.; Ikenaka, T.; Tachida, F.

Cancer Res. 38, 3483-3486, 1978

A/Title: Copper(II)-binding ability of human alpha-fetoprotein.

A/Reference number: A90758; MUID:79001617; PMID:80265

A/Contents: annotation; copper binding

R/Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.

J. Immunol. 143, 1680-1684, 1989

A/Title: Structures of histamine-releasing peptides formed by the action of acid proteases

A/Reference number: A45800; MUID:89341406; PMID:2474609

A/Accession: C45800

A/Status: preliminary

A/Molecule type: protein

A/Residues: 166-173 <CAR>

R/Heard, J.

Mol. Cell. Biol. 7, 2425-2434, 1987

A/Title: Determinants of rat albumin promoter tissue specificity analyzed by an improved

A/Reference number: I57621; MUID:87286876; PMID:3475566

A/Accession: I57621

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-5 <RES>

A/Cross-references: GB:M16825; NID:g202828; PID:AAA40712.1; PID:g554412

C/Superfamily: serum albumin; serum albumin repeat homology

C/Keywords: carrier protein; duplication; metal binding; plasma

F/1-18/Domain: signal sequence; status experimental <SIG>

F/19-24/Domain: propeptide; status experimental <PRO>

F/25-608/Product: serum albumin; status experimental <NAT>

F/29-202/Domain: serum albumin repeat homology <SAI>

F/221-394/Domain: serum albumin repeat homology <SA2>

F/413-592/Domain: serum albumin repeat homology <SA3>

F/27/Binding site: copper (His) #status experimental

F/77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,4

4.4%; Score 26; DB 1; Length 608;

Query Match 100.0%; Pred. No. 1.9e-17; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 ADCCAKQEPNERCEFLQHKDDNP 113

112 ADCCAKQEPNERCEFLQHKDDNP 137

DB

RESULT 7

JC5838

albumin - Mongolian jird

C/Species: Meriones unguiculatus (Mongolian jird)

C/Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 19-May-2000

C/Accession: JC5838

R/Yoshida, K.; Seto-Ohehima, A.; Sinohara, H.

DNA Res. 4, 351-354, 1997

A/Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthesis in the

A/Reference number: JC5838; MUID:98116663; PMID:9455485

A/Accession: JC5838

A/Molecule type: mRNA

A/Residues: 1-609 <YOS>

A/Cross-references: DBJ:AB006197; NID:g2317277; PID:BAA21765.1; PID:g2317278

A/Experimental source: liver

A/Superfamily: serum albumin; serum albumin repeat homology

F/222-395/Domain: serum albumin repeat homology <SA2>

Query Match 4.3%; Score 25; DB 2; Length 609;

Best Local Similarity 100.0%; Pred. No. 1.9e-16;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 EMADCCAKQEPNERCEFLQHKDDNP 110

111 EMADCCAKQEPNERCEFLQHKDDNP 135

DB

RESULT 8

AB05

serum albumin precursor [validated] - bovine

N/Alternate names: 67K protein; prealbumin

C/Species: Bos primigenius taurus (cattle)

C/Date: 24-Apr-1984 #sequence_revision 30-Sep-1993 #text_change 18-Aug-2000

C/Accession: A38895; A36401; A91258; B60808; S10780; A26693; A90309; A91458; A9

R/Holowachuk, E.W.; Stolteneborg, J.K.; Reed, R.G.; Peters Jr., T.

submitted to the EMBL Data Library, August 1991

A/Description: Bovine serum albumin; cDNA sequence and expression.

A/Reference number: A38885

A/Accession: A38885

A/Molecule type: mRNA

A/Residues: 1-607 <HOL>

A/Cross-references: EMBL:M73215

R/Hirayama, K.; Akashi, S.; Furuya, M.; Fukuhata, K.

Biochem. Biophys. Res. Commun. 173, 639-646, 1990

A/Title: Rapid confirmation and revision of the primary structure of bovine serum album

A/Reference number: A36401; MUID:91083649; PMID:2260975

A/Accession: A36401

A/Molecule type: protein

A/Residues: 25-41, 'H', 43-189, 'E', 191-213, 'T', 215-323, 'D', 325-393, 'TS', 396-607 <HIR>

R/MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.

Eur. J. Biochem. 98, 477-485, 1979

A/Title: Biosynthesis of bovine plasma proteins in a cell-free system.

A/Reference number: A91258; MUID:80024278; PMID:488109

A/Accession: A91258

A/Molecule type: protein

A/Residues: 1-32 <MAG>

R/Hsieh, J.C.; Lin, F.P.; Tam, M.F.

Anal. Biochem. 170, 1-8, 1988

A/Title: Electrophoretic transfer of bovine plasma proteins from an analytical isoelectrofocusing

A/Reference number: A60806; MUID:88267456; PMID:3389500

A/Accession: B60808

A/Molecule type: protein

A/Residues: 25-41 <HSI>

R/Strawich, E.; Glimcher, M.J.

Eur. J. Biochem. 191, 47-56, 1990

A/Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelins' is alb

A/Reference number: S10780; MUID:90336641; PMID:2379503

A/Accession: S10780

A/Molecule type: protein

A/Residues: 25-41, 'H', 43-57, 59-64 <STR>

R/Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.

J. Immunol. 143, 1680-1684, 1989

A/Title: Structures of histamine-releasing peptides formed by the action of acid protea

A/Reference number: A45800; MUID:89341406; PMID:2474609

A/Accession: D45800

A/Molecule type: protein

A/Residues: 163-172 <CAR>

R/Carraway, R.E.; Mitra, S.P.; Cochran, D.E.

J. Biol. Chem. 262, 5968-5973, 1987

A/Title: Structure of a biologically active neurotensin-related peptide obtained from p

A/Reference number: A26693; MUID:87194805; PMID:2437111

A/Accession: A26693

A/Molecule type: protein

A/Residues: 165-172, 'L', <CA2>

R/Reed, R.G.; Putnam, F.W.; Peters Jr., T.

Biochem. J. 191, 867-868, 1980

A/Title: Sequence of residues 400-403 of bovine serum albumin.

A/Reference number: A90309; MUID:82023364; PMID:7283978

A/Accession: A90309

A/Molecule type: protein

A/Residues: 402-433 <REE>

R/Brown, J.R.

Fed. Proc. 34, 591, 1975

A/Title: Structure of bovine serum albumin.

A/Reference number: A91458

A/Accession: A91458

A/Molecule type: protein

A;Residues: 25-41,'H',43-117,'EQ',120-179,181-189,'E',191-194,'A',196-213,'T',215-288,'E'
R;Brown, J.R.
Submitted to the Atlas, April 1975
A;Reference number: A94551
A;Accession: A94551
A;Molecule type: protein
A;Residues: 190-195 <BR2>
R;Brown, J.R.
Fed. Proc. 33, 1389, 1974
A;Reference number: A91457
A;Contents: annotation: disulfide bonds
R;Werlen, R.C.; Offord, R.E.; Rose, K.
Biochem. J. 302, 907-911, 1994
A;Title: Preparation and characterization of novel substrates of insulin proteinase (EC
A;Reference number: S55232; MUID:95031935; PMID:7945219
A;Accession: S55232
A;Status: preliminary
A;Molecule type: protein
A;Residues: 529-536;569-572 <WEB>
C;Superfamily: serum albumin, serum albumin repeat homology
C;Keywords: carrier protein; copper binding; duplication; plasma
F;1-18/Domain: signal sequence #status experimental <SIG>
F;19-24/Domain: propeptide #status experimental <PRO>
F;25-607/Product: serum albumin #status experimental <MPT>
F;29-201/Domain: serum albumin repeat homology <SA1>
F;220-393/Domain: serum albumin repeat homology <SA2>
F;412-591/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status predicted
F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4

Query Match 3.8%; Score 22; DB 1; Length 607;
Best Local Similarity 100.0%; Pred. No. 2e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 ECCHGDLLECADRADLAKYIC 265
DB 267 ECCHGDLLECADRADLAKYIC 288
|||||

RESULT 9
ABSHS
serum albumin precursor - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C;Accession: S06936
R;Brown, W.M.; Dziegielewska, K.M.; Foreman, R.C.; Saunders, N.R.
Nucleic Acids Res. 17, 10495, 1989
A;Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.
A;Reference number: S06936; MUID:90098888; PMID:2602160
A;Accession: S06936
A;Molecule type: mRNA
A;Residues: 1-607 <BRO>
A;Cross-references: EMBL:X17055; NID:g1386; PIDN:CAA34903.1; PID:g1387
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the membra
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status predicted <PRO>
F;25-607/Product: serum albumin #status predicted <MAT>
F;29-201/Domain: serum albumin repeat homology <SA1>
F;220-393/Domain: serum albumin repeat homology <SA2>
F;412-591/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status predicted
F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4

Query Match 3.8%; Score 22; DB 1; Length 607;
Best Local Similarity 100.0%; Pred. No. 2e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 ECCHGDLLECADRADLAKYIC 265
DB 267 ECCHGDLLECADRADLAKYIC 288
|||||

Db 267 ECCHGDLLECADRADLAKYIC 288
RESULT 10
I46986
albumin - dog (fragment)
C;Species: Canis lupus familiaris (dog)
C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 20-Aug-1999
C;Accession: I46986
R;Spitzauer, S.; Schweiger, C.; Sperr, W.R.; Pandjaitan, B.; Valent, P.; Muhl, S.; Ebne
J. Allergy Clin. Immunol. 93, 614-627, 1994
A;Title: Molecular characterization of dog albumin as a cross-reactive allergen.
A;Reference number: I46986; MUID:94201492; PMID:7512102
A;Accession: I46986
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-265 <SPI>
A;Cross-references: GB:S72946; NID:G633937; PIDN:AAB30434.1; PID:G633938
C;Superfamily: serum albumin; serum albumin repeat homology
P;7-180/Domain: serum albumin repeat homology <SA2>

Query Match 3.4%; Score 20; DB 2; Length 265;
Best Local Similarity 100.0%; Pred. No. 9.7e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 ECCHGDLLECADRADLAKY 263
DB 54 ECCHGDLLECADRADLAKY 73
|||||

RESULT 11
A05139
serum albumin - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 05-Jun-1987 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
C;Accession: A05139; I48638
R;Winghetti, P.P.; Law, S.W.; Dugaiczkyk, A.
Mol. Biol. Evol. 2, 347-358, 1985
A;Title: The rate of molecular evolution of alpha-fetoprotein approaches that of pseud
A;Reference number: A93055; MUID:88216123; PMID:2452956
A;Accession: A05139
A;Molecule type: mRNA
A;Residues: 1-418 <MIN>
A;Cross-references: GB:M16111; NID:g191764; PIDN:AAA37190.1; PID:g191765
R;Boccaccio, C.; Deschatrette, J.; Meunier-Rotival, M.
Gene 88, 181-186, 1990
A;Title: Empty and occupied insertion site of the truncated LINE-1 repeat located in th
A;Reference number: I48638; MUID:90269606; PMID:1971802
A;Accession: I48638
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 379-453 <BOC>
A;Cross-references: EMBL:X13060; NID:G52939; PIDN:CAA31458.1; PID:G899334
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-104/Domain: serum albumin repeat homology (fragment) <SA1>
F;123-296/Domain: serum albumin repeat homology <SA2>
F;315-453/Domain: serum albumin repeat homology (fragment) <SA3>

Query Match 3.1%; Score 18; DB 2; Length 453;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 KOEPERNECFLOHKDDNP 110
DB 19 KOEPERNECFLOHKDDNP 36
|||||

RESULT 12
B45800
serum albumin - dog (fragment)
C;Species: Canis lupus familiaris (dog)
C;Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 31-Dec-1993

DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S76732
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-289 <KAN>
A;Cross-references: EMBL:D90916; GB:AB001339; NID:g1653715; PIDN:BAAL18644.1; PID:d10193
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 1.4%; Score 8; DB 2; Length 289;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 577 AASQAALG 584
DB 275 AASQAALG 282

RESULT 15
BMBYD5
BUD5 protein - yeast (*Saccharomyces cerevisiae*)
N;Alternate names: protein YCR038c; protein YCR526; protein YCR721
C;Species: *Saccharomyces cerevisiae*
C;Date: 31-Mar-1993 #sequence revision 31-Mar-1993 #text change 16-Jun-2000
C;Accession: S19450; S19752; S26405; A39933; A39934; S12920
R;Dujon, B.; Fairhead, C.; Thierry, A.
submitted to the Protein Sequence Database, March 1992
A;Reference number: S19446
A;Accession: S19450
A;Molecule type: DNA
A;Residues: 1-538 <DUJ>
A;Cross-references: EMBL:X59720; NID:g1907116; PIDN:CAA42305.1; PID:g1907180; GSPDB:GNC
R;Boitout-Fukuhara, M.; Buhler, J.M.; Daignan-Fornier, B.; Doira, C.; Francinques-Gail
submitted to the Protein Sequence Database, March 1992
A;Reference number: S19396
A;Accession: S19752
A;Molecule type: DNA
A;Residues: 1-538 <BO>
A;Cross-references: EMBL:X59720; NID:g1907116; PIDN:CAA42305.1; PID:g1907180; GSPDB:GNC
R;Jaquet, M.; Buhler, J.M.; Iborra, F.; Francinques-Gailard, M.C.; Soustelle, C.
Yeast 7, 881-888, 1991
A;Title: The MAT locus revisited within a 9.8 kb fragment of chromosome III containing
A;Reference number: S26405; MUID:92160397; PMID:1789011
A;Accession: S26405
A;Molecule type: DNA
A;Residues: 1-296, 'L', 298-376, 'S', 378-538 <JAC>
A;Cross-references: EMBL:X63853; NID:g3901; PIDN:CAA45334.1; PID:g3903
R;Chant, J.; Corrado, K.; Pringle, J.R.; Herskowitz, I.
Cell 65, 1213-1224, 1991
A;Title: Yeast BUD5, encoding a putative GDP-GTP exchange factor, is necessary for bud
A;Reference number: A39933; MUID:91292524; PMID:1905981
A;Accession: A39933
A;Molecule type: DNA
A;Residues: 1-5, 'R', 7-120, 'D', 122-296, 'L', 298-376, 'S', 378-538 <CHA>
A;Cross-references: GB:M63552; NID:g171138; PIDN:AAA34460.1; PID:g171139
R;Powers, S.; Gonzales, E.; Christensen, T.; Cubert, J.; Broek, D.
Cell 65, 1225-1231, 1991
A;Title: Functional cloning of BUD5, a CDC25-related gene from *Saccharomyces cerevisiae*
A;Reference number: A39934; MUID:91292525; PMID:1905982
A;Accession: A39934
A;Molecule type: DNA
A;Residues: 1-5, 'R', 8-538 <POW>
A;Cross-references: GB:M66938; NID:g171140; PIDN:AAA34462.1; PID:g171141
R;Thierry, A.; Fairhead, C.; Dujon, B.
Yeast 6, 521-534, 1990
A;Title: The complete sequence of the 8.2 kb segment left of MAT on chromosome III reve
A;Reference number: S12916; MUID:91181345; PMID:1964349
A;Accession: S12920
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 183-538 <THI>

C;Accession: B45800
R;Caraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1580-1584, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid protease
A;Reference number: A45800; MUID:89341406; PMID:2474609
A;Accession: B45800
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-8 <CAR>

Query Match 1.4%; Score 8; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 142 IARRHPYF 149
DB 1 IARRHPYF 8

RESULT 13
D82862
confugal transfer protein XPa0012 [imported] - *Xylella fastidiosa* (strain 9a5c)
C;Species: *Xylella fastidiosa*
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: D82862
R;Anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: D82862
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-273 <SIM>
A;Cross-references: GB:AB003851; NID:g9112238; PIDN:AAF85581.1; GSPDB:GN00130; XFSC:XPao
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Canargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neco, E.; Docena, C.; El-Dorri, H.; Pacinani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laig
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, B
A;Authors: Martins, E.M.F.; Matukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tshuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjowski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
C;Genetics: annotation
A;Contents: annotation
A;Genes: XPa0012
A;Genome: plasmid
A;Note: plasmid pXF5.1
Query Match 1.4%; Score 8; DB 2; Length 273;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 408 LVRYTKKV 415
DB 170 LVRYTKKV 177

RESULT 14
S76732
hypothetical protein - *Synechocystis* sp. (strain PCC 6803)
C;Species: *Synechocystis* sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C;Accession: S76732
R;Xaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

A:Cross-references: EMBL:X56909; NID:g4489; PIDN:CAA40230.1; PID:g4494
C:Genetics:
A:Gene: SGD:BDU5; MIPS:YCR038C
A:Cross-references: SGD:S000634; MIPS:YCR038C
A:Map position: 3R
C:Superfamily: BUD5 protein; CDC25-type guanine nucleotide exchange activator homology
F:304-536/Domain: CDC25-type guanine nucleotide exchange activator homology <SOS>
Query Match 1.4%; Score 8; DB 1; Length 538;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 26 AFAQYLOQ 33
DB 485 AFAQYLOQ 492
RESULT 16
S05532
gamma-glutamyltransferase (EC 2.3.2.2) - pig
N:Alternate names: gamma-glutamyl transpeptidase
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S05532
R:Papadrikopoulou, A.; Frey, A.; Gassen, H.G.
Eur. J. Biochem. 183, 693-698, 1989
A:Title: Cloning and expression of gamma-glutamyl transpeptidase from isolated porcine B
A:Reference number: S05532; MUID:89377838; PMID:2476308
A:Accession: S05532
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-568 <GAP>
A:Cross-references: GB:246922; NID:g600818; PIDN:CAA87031.1; PID:g600819
C:Superfamily: gamma-glutamyltransferase
C:Keywords: aminocyclotransferase; glycoprotein; heterodimer; membrane protein
F:1-379/Product: gamma-glutamyltransferase heavy chain #status predicted <HCH>
F:380-568/Product: gamma-glutamyltransferase light chain #status predicted <LCH>
Query Match 1.4%; Score 8; DB 1; Length 568;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 348 RLAKTYET 355
DB 211 RLAKTYET 218
RESULT 17
WQECPI
phosphotransferase system enzyme I (EC 2.7.3.9) [validated] - Escherichia coli (strain K
C:Species: Escherichia coli
C:Date: 31-Dec-1998 #sequence_revision 31-Dec-1998 #text_change 01-Mar-2002
C:Accession: B29785; B32345; B24035; H28181; G65015
R:Saffen, D.W.; Presper, K.A.; Doering, T.L.; Roseman, S.
J. Biol. Chem. 262, 16241-16253, 1987
A:Title: Sugar transport by the bacterial phosphotransferase system. Molecular cloning a
A:Reference number: A29785; MUID:88058992; PMID:2960675
A:Accession: B29785
A:Molecule type: mRNA
A:Residues: 1-575 <GAP>
A:Cross-references: GB:J02796; NID:g147397; PIDN:AAA24441.1; PID:g147399
R:De Reuse, H.; Danchin, A.
J. Bacteriol. 170, 3827-3837, 1988
A:Title: The ptsH, ptsI, and crr genes of the Escherichia coli phosphoenolpyruvate-depen
A:Reference number: A32345; MUID:88314869; PMID:2457575
A:Accession: B32345
A:Molecule type: DNA
A:Residues: 1-575 <DER1>
A:Cross-references: GB:M11994; NID:g147261; PIDN:AAA24385.1; PID:g147264
R:De Reuse, H.; Roy, A.; Danchin, A.
Gene 35, 199-207, 1985
A:Title: Analysis of the ptsH-ptsI-crr region in Escherichia coli K-12: nucleotide sequ
A:Reference number: A24035; MUID:85286651; PMID:2411636

A:Accession: B24035
A:Molecule type: DNA
A:Residues: 1-56 <DER2>
A:Cross-references: GB:M10425; NID:g147394; PIDN:AAA24439.1; PID:g147396
R:Byrne, C.R.; Monroe, R.S.; Ward, K.A.; Kredich, N.M.
J. Sacteriol. 170, 3150-3157, 1988
A:Title: DNA sequences of the cysK regions of Salmonella typhimurium and Escherichia cc
A:Reference number: A28181; MUID:88257033; PMID:3290198
A:Accession: H28181
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-54 <BYR>
A:Cross-references: GB:M21451; NID:g145684; PIDN:AAA23656.1; PID:g145688
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: G65015
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-575 <BLAT>
A:Cross-references: GB:AB000329; GB:U00096; NID:g2367137; PIDN:AACT5469.1; PID:g1788756
A:Experimental source: strain K-12, substrain MG1655
R:Garrett, D.S.; Gronenborn, A.M.; Clore, G.M.
Submitted to the Brookhaven Protein Data Bank, January 1997
A:Reference number: A68446; PDB:1EZA
A:Contents: annotation; conformation by (1)H-, (13)C-, and (15)N-NMR, residues 1-258, 'R
R:Garrett, D.S.; Seok, Y.J.; Liao, D.I.; Peterkofsky, A.; Gronenborn, A.M.; Clore, G.M
Biochemistry 36, 2517-2530, 1997
A:Title: Solution structure of the 30 kDa N-terminal domain of enzyme I of the Escheric
A:Reference number: A59022; MUID:97207064; PMID:9054557
A:Contents: annotation; conformation by (1)H-, (13)C-, and (15)N-NMR
R:Liao, D.I.; Davies, D.R.
Submitted to the Brookhaven Protein Data Bank, May 1996
A:Reference number: A67861; PDB:1ZYM
A:Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 3-249
R:Liao, D.I.; Silvestro, E.; Seok, Y.J.; Lee, B.R.; Peterkofsky, A.; Davies, D.R.
Structure 4, 861-872, 1996
A:Title: The first step in sugar transport: crystal structure of the amino terminal dom
in HPr.
A:Reference number: A59023; MUID:96434331; PMID:8805571
A:Contents: annotation; X-ray crystallography, 2.5 angstroms
C:Comment: This enzyme acts on the phosphotransferase system phosphohistidine-contains
C:Genetics:
A:Gene: ptsI
A:Map position: 52 min
C:Complex: autophosphorylation requires dimerization
C:Function:
A:Description: catalyzes the phosphorylation of the phosphotransferase system phosphohi
C:Superfamily: phosphotransferase system enzyme I; phosphotransferase system enzyme I h
C:Keywords: homodimer; phosphohistidine; phosphoprotein; phosphotransferase; sugar trar
F:2-561/Domain: phosphotransferase system enzyme I homology <PTI>
F:189/Active site: His (phosphohistidine intermediate) #status experimental
Query Match 1.4%; Score 8; DB 1; Length 575;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 184 ELRDEGKA 191
DB 412 ELRDEGKA 419
RESULT 18
WQEBPI
phosphotransferase system enzyme I (EC 2.7.3.9) - Salmonella typhimurium
N:Contains: phosphoenolpyruvate-protein phosphotransferase (EC 2.7.3.9)
C:Species: Salmonella typhimurium
C:Date: 30-Jun-1990 #sequence_revision 30-Sep-1992 #text_change 08-Sep-2000
C:Accession: A41027; D28181
R:Ucciali, C.; Crocenz, I.S.; Freire, E.; Roseman, S.
J. Biol. Chem. 266, 19519-19527, 1991

A:Title: Sugar transport by the bacterial phosphotransferase system. Structural and the
A:Reference number: A41027; MUID:92011751; PMID:1655788
A:Accession: A41027
A:Molecule type: DNA
A:Residues: 1-575 <LTC>
A:Cross-references: GB:M76176; NID:G153956; PIDN:AAA27060.1; PID:G153957
A:Note: Part of this sequence was confirmed by protein sequencing
R:Byrne, C.R.; Monroe, R.S.; Ward, K.A.; Kredich, N.M.
J. Bacteriol. 170, 3150-3157, 1988
A:Title: DNA sequences of the cysK regions of Salmonella typhimurium and Escherichia coli
A:Reference number: A28181; MUID:88257033; PMID:3290198
A:Accession: D28181
A:Molecule type: DNA
A:Residues: 1-299 <BYR>
A:Cross-references: GB:M76176; NID:G153956; PIDN:AAA27060.1; PID:G153957
C:Comment: The phosphoenolpyruvate-glycose phosphotransferase system (PTS) comprises two
ors (II-A/II-B or III/III), lipid, and divalent cation. The phosphoryl group is transferred
substrate as it crosses the cell membrane.
C:Genetics:
A:Gene: ptsI
A:Map position: 49 min
C:Superfamily: phosphotransferase system enzyme I; phosphotransferase system enzyme I h
C:Keywords: phosphoprotein; phosphotransferase; sugar transport system
F:2-561/Domain: phosphotransferase system enzyme I homology <PT1>
Query Match 1.4%; Score 8; DB 1; Length 575;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;
QY 184 ELRDEGKA 191
DB 412 ELRDEGKA 419
RESULT 19
H91039
PEP:protein phosphotransferase system enzyme I [imported] - Escherichia coli (strain O15
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
A:Accession: H91039
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: H91039
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-575 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA836711.1; PID:G13362758; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECs3288
C:Superfamily: phosphotransferase system enzyme I; phosphotransferase system enzyme I h
Query Match 1.4%; Score 8; DB 2; Length 575;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;
QY 184 ELRDEGKA 191
DB 412 ELRDEGKA 419
RESULT 20
C95884
PEP:protein phosphotransferase system enzyme I [imported] - Escherichia coli (strain O15
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
A:Accession: C95884
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,
Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: C95884
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-575 <STO>
A:Cross-references: GB:AE005174; NID:G12516792; PIDN:AA057535.1; GSPDB:GN00145; UWGP:Z
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: ptsI
C:Superfamily: phosphotransferase system enzyme I; phosphotransferase system enzyme I
Query Match 1.4%; Score 8; DB 2; Length 575;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;
QY 184 ELRDEGKA 191
DB 412 ELRDEGKA 419
RESULT 21
FPHU
alpha-fetoprotein precursor [validated] - human
N:Alternate names: AFP; alpha-1-fetoprotein; alpha-fetoglobulin
C:Species: Homo sapiens (man)
C:Date: 19-Feb-1984 #sequence_revision 19-Feb-1984 #text_change 08-Dec-2000
A:Accession: A26624; S37655; A93961; A91497; A23699; A61480; A90624; A90757; A93042; A
R:Gibbs, P.E.M.; Zielinski, R.; Boyd, C.; Dugaiczkyk, A.
Biochemistry 26, 1332-1343, 1987
A:Title: Structure, polymorphism, and novel repeated DNA elements revealed by a complet
A:Reference number: A26624; MUID:87185438; PMID:2436661
A:Accession: A26624
A:Molecule type: DNA
A:Residues: 1-609 <GIB>
A:Cross-references: GB:M16110; NID:G773678; PIDN:AA58754.1; PID:G178236
R:McVey, J.H.; Michaelides, K.; Hansen, L.P.; Ferguson-Smith, M.; Tilghman, S.; Krumlau
Hum. Mol. Genet. 2, 379-384, 1993
A:Title: A G->A substitution in an HNF 1 binding site in the human alpha-fetoprotein ge
A:Reference number: S37655; MUID:93278385; PMID:7684942
A:Accession: S37655
A:Molecule type: DNA
A:Residues: 1-28 <MCV>
A:Cross-references: EMBL:Z19532; NID:G28527; PIDN:CAA79592.1; PID:G28528
A:Note: The authors translated the codon TAT for residue 26 as Thr
R:Morinaga, T.; Sakai, M.; Wegmann, T.G.; Tamaoki, T.
Proc. Natl. Acad. Sci. U.S.A. 80, 4604-4608, 1983
A:Title: Primary structures of human alpha-fetoprotein and its mRNA.
A:Reference number: A93961; MUID:83273664; PMID:6192439
A:Accession: A93961
A:Molecule type: mRNA
A:Residues: 1-609 <MOR>
A:Cross-references: GB:J00077; NID:G311348; PIDN:CAA24758.1; PID:G31351
R:Beattie, W.G.; Dugaiczkyk, A.
Gene 20, 415-422, 1982
A:Title: Structure and evolution of human alpha-fetoprotein deduced from partial sequen
A:Reference number: A91497; MUID:83158778; PMID:6187626
A:Accession: A91497
A:Molecule type: mRNA
A:Residues: 429-556 <BEA>
A:Cross-references: GB:J00076
R:Pucci, P.; Siciliano, R.; Malorni, A.; Marino, G.; Tecce, M.P.; Ceccarini, C.; Terran
Biochemistry 30, 5061-5066, 1991
A:Title: Human alpha-fetoprotein primary structure: a mass spectrometric study.
A:Reference number: A23699; MUID:91242409; PMID:1709810
A:Accession: A23699
A:Molecule type: protein
A:Residues: 19-45160-97;102-107;122-184;187-249;255-489;507-609 <PUC>
R:Tece, M.F.; Terrana, B.; Gluliani, M.M.; Ceccarini, C.
J. Nucl. Med. Allied Sci. 34, 213-216, 1990
A:Title: Characterization of in vitro expressed human alpha-fetoprotein as highly repro
A:Reference number: A61480; MUID:91235826; PMID:1709209
A:Accession: A61480

A:Molecule type: protein
A:Residues: 19-45163-97;102-107;122-184;187-249;255-489;507-609 <TEC>
R:Yachnin, S.; Hsu, R.; Heinrikson, R.L.; Miller, J.B.
Biochim. Biophys. Acta 493, 418-428, 1977
A:Title: Studies on human alpha-fetoprotein. Isolation and characterization of monomeric
A:Reference number: A90624; MUID:77242506; PMID:70228
A:Accession: A90624
A:Molecule type: protein
A:Residues: S', 20-22, S', 24-35 <AC>
A:Note: dimeric and trimeric forms have been found in addition to the monomeric form
R:Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 37, 3663-3667, 1977
A:Title: Comparative chemical structure of human alpha-fetoproteins from fetal serum and
A:Reference number: A90757; MUID:78001760; PMID:71198
A:Accession: A90757
A:Molecule type: protein
A:Residues: 'S', 20-30, 'A', 32-37, 'A' <AOY>
R:Ruuslahti, E.; Fihko, K.; Vaheri, A.; Seppala, M.; Virolainen, M.; Kontinen, A.
J. Biol. Chem. 260, 5055-5060, 1985
A:Title: The human alpha-fetoprotein gene. Sequence organization and the 5' flanking reg
A:Reference number: A92520; MUID:85182639; PMID:2580830
A:Contents: annotation; gene, exons and introns
R:Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 38, 3483-3486, 1978
A:Title: Copper(II)-binding ability of human alpha-fetoprotein.
A:Reference number: A90758; MUID:79001617; PMID:80265
A:Contents: annotation; metal binding
R:Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 39, 3571-3574, 1979
A:Title: alpha-fetoprotein as a carrier protein in plasma and its bilirubin-binding abil
A:Reference number: A90759; MUID:80001710; PMID:89900
A:Contents: annotation; bilirubin binding
C:Comment: AFP is synthesized by the fetal liver and yolk sac. It occurs in the plasma c
o trace amounts after birth. The serum level in adults is usually less than 40 ng/ml. AFP
C:Comment: Human AFP binds copper, nickel, and fatty acids as well as, and the bilirubin
properties.
C:Genetics:
A:Gene: GDB:AFP
A:Cross-references: GDB:119660; OMIM:104150
A:Map position: 4q11-q13
A:Introns: 29/1, 46/2, 90/3; 161/2, 205/3; 238/2, 281/3; 353/2; 397/3; 430/2; 476/3; 551
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-509/Product: alpha-fetoprotein #status experimental <MAT>
F:29-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:22/Binding site: copper (His) #status experimental
F:29-114,113-124,148-193,192-201,224-270,269-277,289-303,302-313,384-393,416-462,461-472
F:249/Binding site: bilirubin (Lys) #status predicted
F:251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.4%; Score 8; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 YEIARRHP 147
|||||
DB 164 YEIARRHP 171

RESULT 22
FPGO
alpha-fetoprotein precursor - gorilla
C:Species: Gorilla gorilla (gorilla)

```

C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999.
C/Accession: A37970
R/Ryan, S.C.; Zielinski, R.; Dugaiczkyk, A.
Genomics 9, 60-72, 1991
A:/title: Structure of the gorilla alpha-fetoprotein gene and the divergence of primates
A:/reference number: A37970; MUID:91189517; PMID:1706310
A:/Accession: A37970
A:/molecule type: DNA
A:/Residues: 1-609 <RYA>
A:/Cross-references: GB:N38272; NID:g817863; PIDN:AAA73520.1; PID:g177041
C/Genetics:
A:/Map position: 4q11-12
A:/Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 55/3
C:/Superfamily: serum albumin; serum albumin repeat homology
C/Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma
F:/1-18/Domain: signal sequence #status predicted <SIG>
F:/19-609/Product: alpha-fetoprotein #status predicted <MAT>
F:/23-202/Domain: serum albumin repeat homology <SA1>
F:/231-394/Domain: serum albumin repeat homology <SA2>
F:/413-592/Domain: serum albumin repeat homology <SA3>
F:/22/Binding site: copper (His) #status predicted
F:/99-141,113-124,148-193,192-201,224-270,269-277,289-303,302-313,384-393,416-462,461-47
F:/249/Binding site: bilirubin (Iys) #status predicted
F:/251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Watch
Best Local Similarity 1.4%; Score 8; DB 1; Length 609;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 YEIARRHP 147
|||
Db 164 YEIARRHP 171
|||

RESULT 23
JC4258
alpha-fetoprotein precursor - chimpanzee
C/Species: Pan troglodytes (chimpanzee)
C/Date: 27-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 20-Aug-1999
C/Accession: JC4258
R/Nishio, H.; Gibbs, P.E.M.; Minghetti, P.P.; Zielinski, R.; Dugaiczkyk, A.
Gene 162, 213-220, 1995
A:/title: The chimpanzee alpha-fetoprotein-encoding gene shows structural similarity to
A:/reference number: JC4258; MUID:96032345; PMID:7557431
A:/Accession: JC4258
A:/molecule type: DNA
A:/Residues: 1-609 <NIS>
A:/Cross-references: GB:U21916; NID:g841311; PIDN:AAA91641.1; PID:g841312
C/Comment: This protein is a plasma protein produced in the fetal and neonatal liver ar
o similar properties and structure.

```

A: Map position: 3p
A: Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 521/2
C: Superfamily: serum albumin; serum albumin repeat homology
C: Keywords: glycoprotein
F: 1..19/Domain: signal sequence #status predicted <SIG>
F: 20..609/Product: alpha-fetoprotein #status predicted <MAT>
F: 225..202/Domain: serum albumin repeat homology <SA1>
F: 221..394/Domain: serum albumin repeat homology <SA2>
F: 413..592/Domain: serum albumin repeat homology <SA3>
F: 44..251/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

Query Match      1.4%; Score 8; DB 2; Length 609;
Best Local Similarity 100.0%; Freq. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      140 YETARRHP 147
          |||||
          |||||
Db      164 YETARRHP 171

```

RESULT 24

ABCHS
serum albumin precursor - chicken
C:Species: Gallus gallus (Chicken)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: S15571; A05078; A13451
R/Cassidy, A.I.; Salkild, C.K.; Baverstock, P.; Wallace, J.C.
submitted to the EMBL Data Library, July 1991
A:Reference number: S15571
A:Accession: S15571
A:Molecule type: mRNA
A:Residues: 1-615 <CNS>
A:Cross-references: EMBL:X60688; NID:G63747; PIDN:CAA43098.1; PID:G63748
R/Hache, R.J.G.; Wiskocil, R.; Vasa, M.; Roy, R.N.; Lau, P.C.K.; Deeley, R.G.
J. Biol. Chem. 268, 4556-4564, 1993.
A:Title: The 5' noncoding and flanking regions of the avian very low density apolipoprotein
A:Reference number: A05078; MUID:83161037; PMID:6187737
A:Accession: A05078
A:Molecule type: DNA
A:Residues: 1-28 <HAC>
A:Cross-references: GB:V00381; NID:G63038; PIDN:CAA23680.1; PID:G63039
R/Rosen, A.M.; Geller, D.M.
Biochem. Biophys. Res. Commun. 78, 1060-1066, 1977
A:Title: Chicken microsomal albumin: amino terminal sequence of chicken proalbumin.
A:Reference number: A13451; MUID:78019943; PMID:911327
A:Accession: A13451
A:Molecule type: protein
A:Residues: 19-23, 'M', 25-30 <ROS>
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
mones (weak bonds with these hormones promote their transfer across the membranes), thy
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-26/Domain: propeptide #status predicted <PRO>
F:27-613/Product: serum albumin #status predicted <MAT>
F:32-206/Domain: serum albumin repeat homology <SA1>
F:225-398/Domain: serum albumin repeat homology <SA2>
F:417-596/Domain: serum albumin repeat homology <SA3>
F:30/Binding site: copper (His) #status predicted
F:80-89, 102-118, 117-128, 152-197, 196-205, 228-274, 273-281, 293-307, 306-317, 344-389, 388-397,
Query Match 1.4%; Score 8; DB 1; Length 615;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 96 PERNECFL 103
DB 123 PERNECFL 130
RESULT 25
E83610
conserved hypothetical protein PA0285 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Sep-2003
C:Accession: E83610
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: E83610
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-760 <STO>
A:Cross-references: GB:AE004466; GB:AE004091; NID:G9946120; PIDN:AA03674.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0285
C:Superfamily: oxygen sensor diguanylate cyclase/c-di-GMP phosphodiesterase
Query Match 1.4%; Score 8; DB 2; Length 760;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 215 AVARLSQR 222
DB 425 AVARLSQR 432
RESULT 26
C86450
FSD14.27 protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Nov-2001
C:Accession: C86450
R/Rheologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.
Annu. N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, B.; Kim, C
A:Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: C86450
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1090 <STO>
A:Cross-references: GB:AE005172; NID:G9920625; PIDN:AAF81347.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1
Query Match 1.4%; Score 8; DB 2; Length 1090;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 181 KLDELRLDE 188
DB 241 KLDELRLDE 248

RESULT 27
H70178
exodeoxyribonuclease V, beta chain (recB) homolog - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C>Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: H70178
R/Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Whit
son, D.; Peterson, J.; Kexlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: H70178
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1169 <KLE>
A:Cross-references: GB:AE001164; GB:AE000783; NID:G2688550; PIDN:AA06981.1; PID:G26885
A:Experimental source: strain B31
Query Match 1.4%; Score 8; DB 2; Length 1169;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 351 KTYETTLLE 358
DB 684 KTYETTLLE 691

RESULT 28
G36769
17R protein - human herpesvirus 6 (strain Uganda-1102)

C:Species: human herpesvirus 6
 C>Date: 26-Jul-1991 #sequence_revision 26-Jul-1991 #text_change 08-Oct-1999
 C:Accession: G36769
 R:Lawrence, G.L.; Chee, M.; Craxton, M.A.; Gompels, U.A.; Honess, R.W.; Barrell, B.G.
 J. Virol. 64, 287-299, 1990
 A:Title: Human herpesvirus 6 is closely related to human cytomegalovirus.
 A:Reference number: A33560; MUID:90080132; PMID:2152817
 A:Accession: G36769
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-77 <LAW>
 A:Cross-references: GB:M68963; GB:M28243; NID:G325494; PIDN:AAA65579.1; PID:G325511

Query Match 1.2%; Score 7; DB 2; Length 77;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 EKPLEK 286
 |||||
 DB 52 EKPLEK 58

RESULT 29
 C90244
 conserved hypothetical protein [imported] - Sulfolobus solfataricus
 C:Species: Sulfolobus solfataricus
 C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
 C:Accession: C90244
 R:She, Q.; Singh, R.K.; Confalonieri, P.; Zivanovic, Y.; Allard, G.; Awevez, M.J.; Chan-
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
 arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
 submitted to GenBank, April 2001
 A:Description: Sulfolobus solfataricus complete genome.
 A:Reference number: A99139
 A:Accession: C90244
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-122 <KUR>
 A:Cross-references: GB:AE006641; NID:G13814112; PIDN:AAK41210.1; GSPDB:GN00155
 C:Genetics:
 A:Gene: SSO0931

Query Match 1.2%; Score 7; DB 2; Length 122;
 Best Local Similarity 100.0%; Pred. No. 58;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 LVNEVTE 48
 |||||
 DB 73 LVNEVTE 79

RESULT 30
 S41361
 genome polypeptide - hepatitis C virus (genotype 3, N4) (fragment)
 N:Contains: core protein
 C:Species: hepatitis C virus
 A:Variety: genotype 3, N4
 A:Variety: genotype 3, N4
 C:Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
 C:Accession: S41361
 R:van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
 submitted to the EMBL Data Library, January 1994
 A:Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
 A:Reference number: S41341
 A:Accession: S41361
 A:Molecule type: genomic RNA
 A:Residues: 1-123 <VAN>
 A:Cross-references: EMBL:Z29464; NID:G443890; PIDN:CAAG2602.1; PID:G443891
 A:Experimental source: genotype 3, N4
 C:Superfamily: hepatitis C virus genome polypeptide
 C:Keywords: capsid protein; core protein; polypeptide
 P:1-123/Product: core protein #status predicted <MAT>

Query Match 1.2%; Score 7; DB 2; Length 123;

Best Local Similarity 100.0%; Pred. No. 58;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGKV 433
 |||||
 DB 116 SRNLGKV 122

RESULT 31
 S41360
 genome polypeptide - hepatitis C virus (genotype 3, N3) (fragment)
 N:Contains: core protein
 C:Species: hepatitis C virus
 A:Variety: genotype 3, N3
 C:Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
 C:Accession: S41360
 R:van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
 submitted to the EMBL Data Library, January 1994
 A:Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
 A:Reference number: S41341
 A:Accession: S41360
 A:Molecule type: genomic RNA
 A:Residues: 1-124 <VAN>
 A:Cross-references: EMBL:Z29463; NID:G443888; PIDN:CAAG2601.1; PID:G443889
 A:Experimental source: genotype 3, N3
 C:Superfamily: hepatitis C virus genome polypeptide
 C:Keywords: capsid protein; core protein; polypeptide
 P:1-124/Product: core protein #status predicted <MAT>

Query Match 1.2%; Score 7; DB 2; Length 124;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGKV 433
 |||||
 DB 116 SRNLGKV 122

RESULT 32
 S41352
 genome polypeptide - hepatitis C virus (genotype 2, N1) (fragment)
 N:Contains: core protein
 C:Species: hepatitis C virus
 A:Variety: genotype 2, N1
 C:Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
 C:Accession: S41352
 R:van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
 submitted to the EMBL Data Library, January 1994
 A:Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
 A:Reference number: S41341
 A:Accession: S41352
 A:Molecule type: genomic RNA
 A:Residues: 1-125 <VAN>
 A:Cross-references: EMBL:Z29455
 A:Experimental source: genotype 2, N1
 C:Superfamily: hepatitis C virus genome polypeptide
 C:Keywords: capsid protein; core protein; polypeptide
 P:1-125/Product: core protein #status predicted <MAT>

Query Match 1.2%; Score 7; DB 2; Length 125;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGKV 433
 |||||
 DB 116 SRNLGKV 122

RESULT 33
 AC2901
 hypothetical protein sdhD [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
 C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C;Accession: AC2901
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
C;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AC2901
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-126 <KUR>
A;Cross-references: GB:AE008688; PIDN:AAL43625.1; PID:g17741146; GSPDB:GN00186
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: sbhd
A;Map position: circular chromosome

Query Match 1.2%; Score 7; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 AEVSKLV 235
|||||
Db 93 AEVSKLV 99

RESULT 34
E97676
succinate dehydrogenase membrane anchor chain (AF007569) [imported] - Agrobacterium tum
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C;Accession: E97676
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Goldman,
A.; Liu, F.; Wollan, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: E97676
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-126 <KUR>
A;Cross-references: GB:AE007869; PIDN:AAK88366.1; PID:g15157851; GSPDB:GN00169
C;Genetics:
A;Gene: AGR_C 4793
A;Map position: circular chromosome

Query Match 1.2%; Score 7; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 AEVSKLV 235
|||||
Db 93 AEVSKLV 99

RESULT 35
S70820
H+-transporting two-sector ATPase (SC 3.6.3.14) epsilon chain - red alga (Cyanidium cald
C;Species: Chloroplast Cyanidium caldarium
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 03-Jun-2002
C;Accession: S70820; S36413
R;Kostzewa, M.; Zetsche, K.
Plant Mol. Biol. 23, 67-76, 1993
A;Title: Organization of plastid-encoded ATPase genes and flanking regions including hom
A;Reference number: S39512; MUID:194033298; PMID:8219057
A;Accession: S70820
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-138 <EOL>
A;Cross-references: EMBL:X66698; NID:G396522; PIDN:CAA47242.1; PID:g396526
A;Note: the source is designated as Galdieria sulphuraria

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1992
C;Genetics:
A;Gene: atpE
A;Genome: chloroplast
C;Superfamily: H+-transporting ATP synthase epsilon chain
C;Keywords: ATP biosynthesis; chloroplast; hydrolase; membrane-associated complex; thyl

Query Match 1.2%; Score 7; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 291 AEVND 297
|||||
Db 67 AEVND 73

RESULT 36
B83835
Hypothetical protein BH1482 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 03-Jun-2002
C;Accession: B83835
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hir
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: B83835
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-146 <STO>
A;Cross-references: GB:AP001512; GB:BA000004; NID:gi0174030; PIDN:BA05201.1; GSPDB:GN0
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH1482
C;Superfamily: Bacillus subtilis hypothetical protein ykuL

Query Match 1.2%; Score 7; DB 2; Length 146;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 ONLIKON 391
|||||
Db 2 ONLIKON 8

RESULT 37
P82675
transcription regulator MarR/EmrR family XF1490 [imported] - Xylella fastidiosa (strain
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C;Accession: P82675
R;Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seque
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: P82675
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-159 <SIM>
A;Cross-references: GB:AE003978; GB:AE003849; NID:G9106510; PIDN:AAF84299.1; GSPDB:GN00
A;Experimental source: strain 9a5C
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carret,
aa-Neto, E.; Docena, C.; El-Dorzy, H.; Facinani, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.B.; Kuramae, E.E.; Laig
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins,
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E. de Sa, R.G.; Santelli, R.V.; Savasa
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silve

M.; Tshukako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1490

Query Match 1.2%; Score 7; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 516 LSEKERQ 522
DB 132 LSEKERQ 138

RESULT 38
T45261
hypothetical protein [imported] - Methanobacterium thermoautotrophicum (fragment)
C:Species: Methanobacterium thermoautotrophicum
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C:Accession: T45261
R:Harms, U.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z22950
A:Accession: T45261
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-167 <HAR>
A:Cross-references: EMBL:X97589; PIDN:CAA66200.1
A:Experimental source: strain Marburg; DSM 2133

Query Match 1.2%; Score 7; DB 2; Length 167;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 CEKPLLE 285
DB 18 CEKPLLE 24

RESULT 39
B82521
hypothetical protein XF2758 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: B82521
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; PMID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: B82521
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-170 <SIM>
A:Cross-references: GB:AE004081; GB:AE003849; NID:g9107985; PIDN:AAR85543.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Artuda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.B.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Parro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martin, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshukako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2758

Query Match 1.2%; Score 7; DB 2; Length 170;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 VTDLTKV 241
DB 52 VTDLTKV 58

RESULT 40
P80388
Genome polyprotein - hepatitis C virus (isolate GM1) (fragment)
N:Contains: amino end of envelope protein M; carboxyl end of capsid protein C
C:Species: hepatitis C virus
A:Note: host Homo sapiens (man)
C:Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 17-Nov-2000
C:Accession: P80388
R:Fuchs, K.; Motz, M.; Schreier, E.; Zachoval, R.; Deinhardt, F.; Roggendorf, M.
Gene 103, 163-169, 1991
A:Title: Characterization of nucleotide sequences from European hepatitis C virus isola
A:Reference number: JN0265; MUID:91365241; PMID:1653756
A:Accession: P80388
A:Molecule type: genomic RNA
A:Residues: 1-178 <FUC>
A:Cross-references: GB:M61719; NID:g329757; PIDN:AAA45536.1; PID:g387654
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein; envelope protein; polyprotein; transmembrane protein
F:1-109/Product: capsid protein C (fragment) #status predicted <CPC>
F:110-178/Product: envelope protein M (fragment) #status predicted <EPM>
F:163-178/Domain: transmembrane #status predicted <TM>

Query Match 1.2%; Score 7; DB 2; Length 178;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGKV 433
DB 110 SRNLGKV 116

RESULT 41
S32740
polyprotein - hepatitis C virus (isolate Russian) (fragment)
N:Contains: capsid protein C; envelope protein M
C:Species: hepatitis C virus
C:Date: 19-Mar-1997 #sequence_revision 05-Dec-1998 #text_change 17-Nov-2000
C:Accession: S32740
R:Vassiliev, V.B.; Viazov, S.O.; Kotova, E.Y.; Nosikov, V.V.
submitted to the EMBL Data Library, April 1993
A:Description: Evidence of new HCV variant of European isolate in Russia.
A:Reference number: S32740
A:Accession: S32740
A:Molecule type: genomic RNA
A:Residues: 1-189 <VAS>
A:Cross-references: EMBL:X71407
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein; envelope protein; polyprotein
F:1-115/Product: capsid protein C #status predicted <CPC>
F:116-189/Product: envelope protein M #status predicted <EPM>

Query Match 1.2%; Score 7; DB 2; Length 189;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGKV 433
DB 116 SRNLGKV 122

RESULT 42
C65266
NAD(P)H-flavin oxidoreductase homolog - Archaeoglobus fulgidus

C;Species: Archaeoglobus fulgidus
 C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
 C;Accession: C59266
 R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 R;Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 R;Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997

A;Authors: Uutterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
 A;Reference number: A69250; MUID:98049343; PMID:9389475

A;Accession: C69266
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA

A;Residues: 1-194 <KLE>
 A;Cross-references: GB:AE001097; GB:AE000782; NID:92689420; PIDN:AAB91098.1; PID:9265051
 C;Superfamily: Methanobacterium thermoautotrophicum NADPH-oxidoreductase

Query Match 1.2%; Score 7; DB 2; Length 194;
 Best Local Similarity 100.0%; Pred. No. 87;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 571 EGKQLVA 577
 |||||
 Db 160 EGKQLVA 166

RESULT 43

E82910
 hypothetical protein U0287 [imported] - Ureaplasma urealyticum

C;Species: Ureaplasma urealyticum
 C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 C;Accession: E82910
 R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
 submitted to GenBank, February 2000

A;Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
 A;Reference number: A82870

A;Accession: E82910
 A;Status: preliminary
 A;Molecule type: DNA

A;Residues: 1-201 <GLA>
 A;Cross-references: GB:AE002125; GB:AF222894; NID:96899253; PIDN:AAF30696.1; GSPDB:GN001
 A;Experimental source: serovar 3; biovar 1

C;Genetics:
 A;Gene: U0287
 A;Genetic code: SGC3

Query Match 1.2%; Score 7; DB 2; Length 201;
 Best Local Similarity 100.0%; Pred. No. 90;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 EVAHREFK 12
 |||||
 Db 91 EVAHREFK 97

RESULT 44

A07063
 hypothetical protein Rv1556 - Mycobacterium tuberculosis (strain H37RV)

C;Species: Mycobacterium tuberculosis
 C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C;Accession: A07063

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
 Connor, R.; Davies, R.; Devlin, K.; Beldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998

A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Accession: A07063
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA

A;Residues: 1-202 <COL>

A;Cross-references: GB:Z74020; GB:AL123456; NID:93261584; PIDN:CAA98307.1; PID:g1403486
 A;Experimental source: strain H37RV

C;Genetics:
 A;Gene: Rv1556
 C;Superfamily: Bacillus subtilis probable transcription regulator yrhI

Query Match 1.2%; Score 7; DB 2; Length 202;
 Best Local Similarity 100.0%; Pred. No. 90;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 423 LVEVSRN 429
 |||||
 Db 114 LVEVSRN 120

RESULT 45

C71691

ribosomal protein S4 - Rickettsia prowazekii

C;Species: Rickettsia prowazekii
 C;Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
 C;Accession: C71691

R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alemark,
 Nature 396, 133-140, 1998

A;Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
 A;Reference number: A71630; MUID:99039499; PMID:9823893

A;Accession: C71691
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA

A;Residues: 1-205 <AND>
 A;Cross-references: GB:AJ235271; GB:AJ235269; NID:93866717; PIDN:CAA14805.1; PID:g38609
 A;Experimental source: strain Madrid E

C;Genetics:
 A;Gene: rpsD; RP345
 C;Superfamily: Escherichia coli ribosomal protein S4

Query Match 1.2%; Score 7; DB 2; Length 205;
 Best Local Similarity 100.0%; Pred. No. 91;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 194 AKORLKC 200
 |||||
 Db 54 AKORLKC 60

RESULT 46

D97758

308 ribosomal protein S4 [imported] - Rickettsia conorii (strain Malish 7)

C;Species: Rickettsia conorii
 C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001
 C;Accession: D97758

R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R
 Science 293, 2093-2098, 2001

A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
 A;Reference number: A97700; MUID:21442074; PMID:11557893

A;Accession: D97758
 A;Status: preliminary
 A;Molecule type: DNA

A;Residues: 1-205 <KUR>
 A;Cross-references: GB:AE006914; PIDN:AAL03006.1; PID:g15619541; GSPDB:GN00173

C;Genetics:
 A;Gene: rpsD
 C;Superfamily: Escherichia coli ribosomal protein S4

Query Match 1.2%; Score 7; DB 2; Length 205;
 Best Local Similarity 100.0%; Pred. No. 91;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 194 AKORLKC 200
 |||||
 Db 54 AKORLKC 60

RESULT 47

T46434
 Hypothetical protein DKFZp344H226.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
 C:Accession: T46434
 R:Ascorge, W.; Winkler, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: Z33028
 A:Accession: T46434
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-205 <AAA>
 A:Cross-references: EMBL:AL137370
 A:Experimental source: adult testis; clone DKFZp344H226
 C:Genetics:
 A:Gene: xkdn
 A>Note: DKFZp344H226.1

Query Match 1.2%; Score 7; DB 2; Length 205;
 Best Local Similarity 100.0%; Pred. No. 91;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 502 FNAETFT 508
 |||||
 DB 170 FNAETFT 176

RESULT 48
 F70708
 Hypothetical protein Rv0779c - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 28-Jul-2000
 C:Accession: F70708
 R:Cole, S.T.; Broech, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: F70708
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-206 <COL>
 A:Cross-references: GB:280226; GB:AL123456; NID:93261638; PIDN:CAB02391.1; PID:e266569;
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: Rv0779c
 C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0779c

Query Match 1.2%; Score 7; DB 2; Length 206;
 Best Local Similarity 100.0%; Pred. No. 92;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 577 AASQAAL 583
 |||||
 DB 87 AASQAAL 93

RESULT 49
 E59732
 PBX3 prophage ORF xkdn - Bacillus subtilis
 C:Species: Bacillus subtilis
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
 C:Accession: E59732
 R:Xunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Arevedo, V.; Berte, C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Title: Fungal, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.P.; Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauesel, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serc, A.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, T.; Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Accession: E59732
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-208 <KUN>
 A:Cross-references: GB:Z99110; GB:AL009126; NID:92633472; PIDN:CAB13124.1; PID:e1183287
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: xkdn

Query Match 1.2%; Score 7; DB 2; Length 208;
 Best Local Similarity 100.0%; Pred. No. 92;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 570 EEGKKLV 576
 |||||
 DB 144 EEGKKLV 150

RESULT 50
 E69020
 conserved hypothetical protein MTH1152 - Methanobacterium thermoautotrophicum (strain D
 C:Species: Methanobacterium thermoautotrophicum
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C:Accession: E69020
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafora, R.; Vicare, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.; K.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
 A:Reference number: A69000; MUID:98037514; PMID:9371463
 A:Accession: E69020
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-213 <MTH>
 A:Cross-references: GB:AE000885; GB:AE000656; NID:92622256; PIDN:AA85641.1; PID:g26222
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH1152
 A:Start codon: TTG

Query Match 1.2%; Score 7; DB 2; Length 213;
 Best Local Similarity 100.0%; Pred. No. 94;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 279 CEKPLE 285
 |||||
 DB 21 CEKPLE 27

RESULT 51
 S43591
 M04D8.4 protein - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
 C:Accession: S43591
 R:Percy, C.; Mortimore, B.
 submitted to the EMBL Data Library, April 1994
 A:Reference number: S43588
 A:Accession: S43591
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <PER>
 A:Cross-references: EMBL:Z32682; NID:9474797; PID:g474801
 C:Genetics:
 A:Introns: 60/1; 182/1
 C:Superfamily: Caenorhabditis elegans M04D8.4 protein

Query Match 1.2%; Score 7; DB 2; Length 218;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 468 PVSDRV 474
|||||
80 PVSDRV 86

Db

RESULT 52
S77870
N:Alternate names: protein MC152
C:Species: Mycoplasma capricolum
C:Date: 09-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 07-Dec-1999
A:Accession: S77870
A:Reference number: S77870
A:Experimental source: ATCC 27343
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C:Genetics:
A:Gene: dnaK
A:Genetic code: SGC3
C:Function:
A:Description: involved in protein folding and assembling/disassembling of protein complex
C:Superfamily: heat shock protein 70
C:Keywords: ATP; molecular chaperone

Query Match 1.2%; Score 7; DB 2; Length 227;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 317 KNVAEAK 323
|||||
100 KNVAEAK 106

Db

RESULT 53
T29034
N:Alternate names: protein MC152
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
A:Accession: T29034
A:Reference number: T29034
A:Experimental source: strain Bristol N2; clone F53G12
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Residues: 1-244 <WUX>
A:Cross-references: EMBL:AF003139; PIDN:AA854165.1; GSDB:GN000019; CESP:F53G12.10
A:Experimental source: strain Bristol N2; clone F53G12
C:Genetics:
A:Gene: CESP:F53G12.10
A:Map position: 1
A:Introns: 4/2; 37/3; 103/3; 176/1
C:Superfamily: rat ribosomal protein L7

Query Match 1.2%; Score 7; DB 2; Length 244;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 412 TKKVPQV 418
|||||

Db

RESULT 54
S64849
N:Alternate names: hypothetical protein Li722
C:Species: Saccharomyces cerevisiae
C:Date: 01-Aug-1995 #sequence_revision 24-May-1996 #text_change 19-Apr-2002
A:Accession: S64849; S64844
A:Reference number: S64845
A:Experimental source: strain S288C
A:Residues: 1-250 <BVR>
A:Cross-references: EMBL:Z33106; NID:G516183; PIDN:CAA93764.1; PID:G530439
A:Experimental source: ATCC 27343
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C:Genetics:
A:Gene: dnaK
A:Genetic code: SGC3
C:Function:
A:Description: involved in protein folding and assembling/disassembling of protein complex
C:Superfamily: heat shock protein 70
C:Keywords: ATP; molecular chaperone

Query Match 1.2%; Score 7; DB 2; Length 250;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 516 LSEKERQ 522
|||||
95 LSEKERQ 101

Db

RESULT 55
PQ0393
N:Contains: capsid protein C; envelope protein M; nonstructural protein NS3; nonstructural protein NS4; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 17-Nov-2000
A:Accession: PQ0393; PQ0398; PQ0399
A:Reference number: PQ0393; PQ0398; PQ0399
A:Experimental source: strain S288C
A:Residues: 135-250 <VAN>
A:Cross-references: EMBL:Z33106; NID:G516183; PIDN:CAA93764.1; PID:G530439
A:Experimental source: ATCC 27343
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C:Genetics:
A:Gene: dnaK
A:Genetic code: SGC3
C:Function:
A:Description: involved in protein folding and assembling/disassembling of protein complex
C:Superfamily: heat shock protein 70
C:Keywords: ATP; molecular chaperone

Query Match 1.2%; Score 7; DB 2; Length 266;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGV 433
|||||
112 SRNLGV 118

Db

RESULT 56
AG1037

[illegible]

Db 4 TKKVPQV 10

RESULT 54
S64849
Hypothetical protein YLR022c - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: hypothetical protein Li722
C:Species: *Saccharomyces cerevisiae*
C>Date: 01-Aug-1995 #sequence_revision 24-May-1996 #text_change 19-Apr-2002
C:Accession: S64849; S64844
R:Obermaier, B.; Piravandi, E.; Rinke, M.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64845
A:Accession: S64849
A:Molecule type: DNA
A:Residues: 1-250 <DB>
A:Cross-references: EMBL:Z73194; NID:gi360327; PIDN:CAA97545.1; PID:e245773; PID:gl3603
A:Experimental source: strain S288C
R:vandenbol, M.; Portetelle, D.; Hilger, F.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64742
A:Accession: S64844
A:Molecule type: DNA
A:Residues: 135-250 <VAN>
A:Cross-references: EMBL:Z73194; MIPS:YLRO22C
A:Experimental source: strain S288C
C:Genetics:
A:Cross-references: SGD:S0004012
A:Map position: 12R

Query Match 1.2%; Score 7; DB 2; Length 250;
Best Local Similarity 100.0%; Pred.No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 516 LSEKERQ 522
|||||

Db 95 LSEKERQ 101
|||||

RESULT 55
PQ0393
genome polyprotein - hepatitis C virus (isolate E-bl) (fragments)
N:Contains: capsid protein C; envelope protein M; nonstructural protein NS3; nonstructu
C:Species: hepatitis C virus
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 17-Nov-2000
C:Accession: PQ0393; PQ0398; PQ0399
J.Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, P
U. Gen. Virol. 73, 1131-1141, 1992
A>Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to
A:Reference number: PQ0393; UID:92268871; PMID:1316939
A:Accession: PQ0393
A:Molecule type: genomic RNA
A:Residues: 1-266 <CHA>
A:Cross-references: DDJ:P10123
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein; envelope protein; nonstructural protein; polyprotein
F:1-111/Product: capsid protein C (fragment) #status predicted <CP>
F:112-124/Product: envelope protein M (fragment) #status predicted <EPM>
F:125-163/Product: nonstructural protein NS3 (fragment) #status predicted <NS3>
F:164-181/Product: nonstructural protein NS4a (fragment) #status predicted <NS4>
F:182-266/Product: nonstructural protein NS5 (fragment) #status predicted <NS5>

Query Match 1.2%; Score 7; DB 2; Length 266;
Best Local Similarity 100.0%; Pred.No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGV 433
|||||

Db 112 SRNLGV 118
|||||

RESULT 56
AG1037

probable capsid protein STV4626 [imported] - Salmonella enterica subsp. enterica serovar C; Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: This species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AG1037
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moutle, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar C; Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AG1037
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-277 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD06748.1; PID:G16505399; GSPDB:GN00176
 C:Genetics:
 A:Gene: STV4626

Query Match 1.2%; Score 7; DB 2; Length 277;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 528 ALVELVK 534
 |||||
 Db 91 ALVELVK 97

RESULT 57
 S30124
 hypothetical protein 3 - Streptomyces cacaoi (fragment)
 C:Species: Streptomyces cacaoi
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 22-Oct-1999
 C:Accession: S30124
 R:Benini, V.M.; Magdalena, J.; Fraipont, C.; Joris, B.; Matagne, A.; Dusat, J.
 Mol. Gen. Genet. 235, 41-48, 1992
 A:Title: Induction of a Streptomyces cacaoi beta-lactamase gene cloned in S. lividans.
 A:Reference number: S30122; MUID:93062806; PMID:1435729
 A:Accession: S30124
 A:Molecule type: DNA
 A:Residues: 1-278 <LEN>
 A:Cross-references: EMBL:X63780; NID:G46880; PIDN:CAA45316.1; PID:G581628
 C:Genetics:
 A:Start codon: GTG

Query Match 1.2%; Score 7; DB 2; Length 278;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 344 VLLRLA 350
 |||||
 Db 101 VLLRLA 107

RESULT 58
 T19411
 hypothetical protein C23H4.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T19411
 R:Wilkinson, J.
 submitted to the EMBL Data Library, August 1996
 A:Reference number: Z19121
 A:Accession: T19411
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-283 <WIL>
 A:Cross-references: EMBL:Z78416; PIDN:CAB01678.1; GSPDB:GN000028; CESP:C23H4.4
 C:Genetics:
 A:Gene: CESP:C23H4.4
 A:Map position: X

A:Introns: 64/3; 97/2; 130/3; 198/1

Query Match 1.2%; Score 7; DB 2; Length 283;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 YEIARRH 146

|||||

Db 271 YEIARRH 277

RESULT 59

B81404
 probable transcription regulator Cj0571 [imported] - Campylobacter jejuni (strain NCTC 81404)
 C:Species: Campylobacter jejuni
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
 C:Accession: B81404
 R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilli, C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barre, Nature 403, 665-668, 2000
 A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy
 A:Reference number: A81250; MUID:20150912; PMID:10688204
 A:Accession: B81404
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-290 <PAR>
 A:Cross-references: GB:AL111168; GB:AL111168; NID:G6967817; PIDN:CAB75207.1; PID:G69680
 A:Experimental source: serotype O2, strain NCTC 11166
 C:Genetics:
 A:Gene: Cj0571

Query Match 1.2%; Score 7; DB 2; Length 290;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 518 EKERQIK 524

|||||

Db 142 EKERQIK 148

RESULT 60

T06510
 probable glutathione transferase (EC 2.5.1.18) gsta2 - wheat
 C:Species: Triticum aestivum (common wheat)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
 C:Accession: T06510
 R:Mauch, F.; Hettig, C.; Rehmann, G.; Bull, J.; Dudler, R.
 Plant Mol. Biol. 16, 1083-1091, 1991
 A:Title: A wheat glutathione-S-transferase gene with transposon-like sequences in the
 A:Reference number: Z15725; MUID:91322503; PMID:1650615
 A:Accession: T06510
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-291 <MAU>
 A:Cross-references: EMBL:X56004; NID:G21789; PIDN:CAA39480.1; PID:G21790
 A:Experimental source: cv. Cheyenne
 C:Genetics:
 A:Gene: gsta2
 A:Introns: 50/3; 67/1
 C:Superfamily: glutathione transferase
 C:Keywords: transferase

Query Match 1.2%; Score 7; DB 2; Length 291;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 370 YAKVFDE 376

|||||

Db 189 YAKVFDE 195

RESULT 61

B71042

hypothetical protein PH1626 - Pyrococcus horikoshii
 C/Species: Pyrococcus horikoshii
 C/Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
 C/Accession: B71042
 R/Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
 DNA Res. 5, 55-76, 1998
 A/Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
 A/Reference number: A71000; UID:98344137; PMID:9679194
 A/Accession: B71042
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-294 <RAW>
 A/Cross-references: GB:AP000006; NID:G3236133; PIDN:BAA30738.1; PID:G3258055
 A/Experimental source: strain OT3
 A/Note: this accession replaces an interim accession for a sequence replaced by GenBank
 C/Genetics:
 A/Genes: PH1626
 C/Superfamily: conserved hypothetical protein YR008w

Query Match 1.2%; Score 7; DB 2; Length 294;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 FKDLGEE 17
 |||||
 DB 29 FKDLGEE 35

RESULT 62
 AC0258
 Probable phage protein YP02116 [imported] - Yersinia pestis (strain CO92)
 C/Species: Yersinia pestis
 C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
 C/Accession: AC0258
 R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
 Nature 413, 523-527, 2001
 A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A/Reference number: AB0001; UID:21470413; PMID:11586360
 A/Accession: AC0258
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-306 <KUR>
 A/Cross-references: GB:AL590842; PIDN:CAC90927.1; PID:G15980123; GSPDB:GN00175
 C/Genetics:
 A/Genes: YP02116

Query Match 1.2%; Score 7; DB 2; Length 306;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 349 LAKTYET 355
 |||||
 DB 13 LAKTYET 19

RESULT 63
 AE3585
 ferric anguibactin transport system permease protein fatD [imported] - Brucella melitens
 C/Species: Brucella melitensis
 C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
 C/Accession: AE3585
 R/DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
 Mazur, M.; Goltzman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A/Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
 A/Reference number: AD3252; PMID:11756688
 A/Accession: AE3585
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-308 <KUR>

A/Cross-references: GB:AE008918; PIDN:AAL53848.1; PID:G17984784; GSPDB:GN00191
 A/Experimental source: strain 16M
 C/Genetics:
 A/Genes: BMEI0066
 A/Map position: II
 C/Superfamily: vitamin B12 transport protein btuc

Query Match 1.2%; Score 7; DB 2; Length 308;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 SLHTLFG 71
 |||||
 DB 15 SLHTLFG 21

RESULT 64
 JN0265
 genome polyprotein - hepatitis C virus (isolate GM2) (fragments)
 N/Contains: amino end of envelope protein M; carboxyl end of capsid protein C; fragment
 C/Species: hepatitis C virus
 A/Note: host Homo sapiens (man)
 C/Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 17-Nov-2000
 C/Accession: JN0265
 R/Fuchs, K.; Mott, M.; Schreier, E.; Zachoval, R.; Deinhardt, F.; Roggendorf, M.
 Gene 103, 163-169, 1991
 A/Title: Characterization of nucleotide sequences from European hepatitis C virus isolat
 A/Reference number: JN0265; UID:91365241; PMID:1653756
 A/Accession: JN0265
 A/Molecule type: genomic RNA
 A/Residues: 1-322 <FUC>
 A/Cross-references: GB:M61717; GB:M61718
 A/Note: the authors translated the codon ACA for residue 198 as Tyr
 C/Superfamily: hepatitis C virus genome polyprotein
 C/Keywords: capsid protein; envelope protein; glycoprotein; polyprotein; transmembrane
 F/1-109/Product: capsid protein C (fragment) #status predicted <COR>
 F/110-178/Product: envelope protein M (fragment) #status predicted <EPM>
 F/163-178/Domain: transmembrane #status predicted <TM1>
 F/179-322/Product: major envelope protein E (fragment) #status predicted <ENV>
 F/253-269/Domain: transmembrane #status predicted <TM2>
 F/191,216/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.2%; Score 7; DB 2; Length 322;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGKV 433
 |||||
 DB 110 SRNLGKV 116

RESULT 65
 A54932
 zeta-crystallin / quinone reductase (NADPH) (EC 1.6.-.-) - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 25-Apr-1995 #sequence_revision 07-Jul-1995 #text_change 11-Jun-1999
 C/Accession: A54932; S42273
 R/Gonzalez, P.; Hernandez-Calzadilla, C.; Rao, P.V.; Rodriguez, I.R.; Zigler Jr., J.S.,
 Mol. Biol. Evol. 11, 305-315, 1994
 A/Title: Comparative analysis of the zeta-crystallin/quinone reductase gene in guinea p
 A/Reference number: A54932; UID:94224126; PMID:8170370
 A/Accession: A54932
 A/Molecule type: mRNA
 A/Residues: 1-331 <GON>
 A/Cross-references: GB:S70056; NID:G546493; PIDN:AAB30620.1; PID:G546494
 A/Experimental source: liver
 A/Note: sequence extracted from NCBI backbone (NCBI:147625, NCBI:147626)
 A/Note: translation of initiator Met is not shown
 R/Josernvall, H.; Persson, B.; du Bois, G.C.; Lavers, G.C.; Chen, J.H.; Gonzalez, P.; Ra
 FEBS Lett. 322, 240-244, 1993
 A/Title: zeta-Crystallin versus other members of the alcohol dehydrogenase super-family
 A/Reference number: S42272; UID:93252077; PMID:8486156
 A/Contents: annotation

Science 273, 813-816, 1996

A:Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re

A:Reference number: 220876; MUID:96325459; PMID:8670425

A:Accession: T30754

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-354 <SEN>

A:Cross-references: EMBL:U60315; NID:g1491943; PIDN:AAC55280.1; PID:g1492095

C:Genetics:

A:Note: MC152R

C:Superfamily: 3beta-hydroxy-Delta5-steroid dehydrogenase

C:Keywords: intramolecular oxidoreductase; isomerase; NAD; oxidoreductase; steroid biosy

Query Match 1.2%; Score 7; DB 2; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 345 LLLRLAK 351
|||||
Db 189 LLLRLAK 195

RESULT 71

A48931

transmembrane glycoprotein CD68, 110K - human

C:Species: Homo sapiens (man)

C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 17-Mar-2000

C:Accession: A48931

R:Holness, C.L.; Simmons, D.L.

Blood 81, 1507-1613, 1993

A:Title: Molecular cloning of CD68, a human macrophage marker related to lysosomal glyco

A:Reference number: A48931; MUID:93200523; PMID:7690921

A:Accession: A48931

A>Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-354 <HO>

A:Cross-references: GB:S57235; NID:g298664; PIDN:AAB25811.1; PID:g298665

A:Experimental source: promonocyte cell line U937

A:Note: sequence extracted from NCBI Backbone (NCBIN:127492, NCBIP:127493)

C:Superfamily: lysosome-associated membrane protein

C:Keywords: glycoprotein; transmembrane protein

Query Match 1.2%; Score 7; DB 2; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 ALVLIAP 27
|||||
Db 335 ALVLIAP 341

RESULT 72

AE3175

hypothetical protein attC [imported] - Agrobacterium tumefaciens (strain C58, Dupont) pl

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C:Accession: AE3175

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AE3175

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-360 <KUR>

A:Cross-references: GB:AEQ08687; PIDN:AAL45819.1; PID:g17743558; GSPDB:GN00188

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: attC

Query Match 1.2%; Score 7; DB 2; Length 360;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 ALVLIAP 27
|||||
Db 335 ALVLIAP 341

RESULT 73

C72590

probable hexosyltransferase (EC 2.4.1.1-) APE1191 [similarity] - Aeropyrum pernix (strai

N:Alternate names: probable capM protein

C:Species: Aeropyrum pernix

C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 16-Jun-2000

C:Accession: C72590

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Taka

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: C72590

A:Molecule type: DNA

A:Residues: 1-363 <YAW>

A:Cross-references: DDBJ:AP000061; NID:g5104821; PIDN:BAA80177.1; PID:g5104863

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE1191

C:Superfamily: probable hexosyltransferase ytxN

C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 1.2%; Score 7; DB 1; Length 363;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 LYEIARR 145
|||||
Db 23 LYEIARR 29

RESULT 74

S43159

outer membrane porin ompF precursor - Salmonella typhimurium

C:Species: Salmonella typhimurium

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Aug-1999

C:Accession: S43159

R:Venegas, A.; Gomez, I.; Bruce, E.; Martinez, M.

submitted to the EMBL Data Library, March 1994

A:Description: PCR amplification and cloning of the Salmonella typhimurium ompF porin 9

A:Reference number: S43159

A:Accession: S43159

A:Molecule type: DNA

A:Residues: 1-363 <VEN>

A:Cross-references: EMBL:Z31594; NID:g469740; PIDN:CAA83471.1; PID:g468741

C:Genetics:

A:Gene: ompF

C:Superfamily: outer membrane protein phoE

C:Keywords: membrane protein; porin; trimer

F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-363/Product: outer membrane porin ompF #status predicted <MAT>

Query Match 1.2%; Score 7; DB 2; Length 363;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 258 ADLAKYI 264
|||||
Db 310 ADLAKYI 316

RESULT 75

AE0616
outer membrane protein F precursor STY1002 [imported] - Salmonella enterica subsp. enterica
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AE0616
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AE0616
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-363 <PAR>
A:Cross-references: GS:AL513382; PIDN:CAD05399.1; PID:gl6502161; GSPDB:GN00176
C:Genetics:
A:Gene: STY1002
C:Superfamily: outer membrane protein phoE

Query Match 1.2%; Score 7; DB 2; Length 363;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 258 ADLAKYI 264
|||
Db 310 ADLAKYI 316

Search completed: April 19, 2004, 16:07:50
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 16:02:34 ; Search time 18 Seconds
(without alignments)
1692.280 Million cell updates/sec

Title: US-09-832-929-18

Perfect score: 585

Sequence: 1 DAHSEVAHRFKDGEENFK.....TCFAEKGKLVASQAALGL 585

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	585	100.0	609	1	ALBU_HUMAN
2	75	12.8	600	1	ALBU_MACMU
3	29	5.0	608	1	ALBU_FELCA
4	27	4.6	607	1	ALBU_HORSE
5	26	4.4	605	1	ALBU_PIG
6	26	4.4	608	1	ALBU_RAT
7	25	4.3	609	1	ALBU_MERIN
8	23	3.9	608	1	ALBU_RABIT
9	22	3.8	607	1	ALBU_BOVIN
10	22	3.8	608	1	ALBU_SHEEP
11	22	3.8	608	1	ALBU_CANFA
12	18	3.1	608	1	ALBU_MOUSE
13	10	1.7	609	1	FETA_HORSE
14	8	1.4	9	1	NEUX_HUMAN
15	8	1.4	457	1	IRF7_MOUSE
16	8	1.4	503	1	IRF7_HUMAN
17	8	1.4	538	1	BUD5_YEAST
18	8	1.4	568	1	GGF1_PIG
19	8	1.4	575	1	PTI_ECOLI
20	8	1.4	575	1	PTI_SALTY
21	8	1.4	609	1	FETA_GORGO
22	8	1.4	609	1	FETA_HUMAN
23	8	1.4	609	1	FETA_PANTR
24	8	1.4	615	1	ALBU_CHICK
25	8	1.4	1169	1	EX5B_BORBU
26	7	1.2	138	1	ULI1_HSVSU
27	7	1.2	138	1	ATPE_GALSI
28	7	1.2	138	1	EX44_PIG
29	7	1.2	149	1	NPC2_MOUSE
30	7	1.2	202	1	YF56_MYCTU
31	7	1.2	205	1	RS4_FICCN
32	7	1.2	205	1	RS4_RICPR
33	7	1.2	208	1	XKDN_BACSU

34	7	1.2	224	1	RNH2_SYNPX
35	7	1.2	243	1	RECO_VIBPA
36	7	1.2	244	1	RL7_CAEEL
37	7	1.2	250	1	YL22_YEAST
38	7	1.2	259	1	PR12_YEAST
39	7	1.2	251	1	GTH2_TORAC
40	7	1.2	251	1	GTH2_WHEAT
41	7	1.2	251	1	YG26_PYRHO
42	7	1.2	328	1	ANR2_MOUSE
43	7	1.2	331	1	QOR_MOUSE
44	7	1.2	332	1	CX14_HUMAN
45	7	1.2	337	1	RUVB_CHLCV
46	7	1.2	339	1	ADD_STRVG
47	7	1.2	348	1	RECA_ENTPA
48	7	1.2	354	1	CD88_HUMAN
49	7	1.2	363	1	OMPF_SALTI
50	7	1.2	363	1	OMPF_SALTY
51	7	1.2	364	1	AROB_CANBF
52	7	1.2	377	1	PYRC_ARATH
53	7	1.2	394	1	CEGT_HUMAN
54	7	1.2	394	1	CEGT_MOUSE
55	7	1.2	394	1	CEGT_RAT
56	7	1.2	399	1	PQOE_BRAJA
57	7	1.2	401	1	RENS_MOUSE
58	7	1.2	405	1	ARRS_CANFA
59	7	1.2	410	1	AUP1_MOUSE
60	7	1.2	417	1	HS47_MOUSE
61	7	1.2	418	1	PROA_CLOAB
62	7	1.2	419	1	P47K_PSECL
63	7	1.2	476	1	AUP1_HUMAN
64	7	1.2	500	1	TRPE_RHOSH
65	7	1.2	513	1	POLG_HCVJ2
66	7	1.2	520	1	POLG_HCVH4
67	7	1.2	520	1	POLG_HCVHK
68	7	1.2	523	1	C9B1_GLYEC
69	7	1.2	586	1	HO_YEAST
70	7	1.2	591	1	DNAK_MYCCA
71	7	1.2	599	1	AFAM_HUMAN
72	7	1.2	605	1	NKX4_HUMAN
73	7	1.2	605	1	NKX4_MOUSE
74	7	1.2	608	1	AFAM_RAT
75	7	1.2	611	1	AFAM_MOUSE
76	7	1.2	617	1	VP3B_HUMAN
77	7	1.2	617	1	VP3B_MOUSE
78	7	1.2	617	1	VP3B_RAT
79	7	1.2	636	1	PET3_YEAST
80	7	1.2	671	1	DNJJ_SCOLI
81	7	1.2	731	1	GLGB_COREF
82	7	1.2	737	1	POLG_HCVJ7
83	7	1.2	773	1	PIGR_RABIT
84	7	1.2	874	1	SYA_HAEIN
85	7	1.2	874	1	SYA_PASMU
86	7	1.2	874	1	SYA_SHEON
87	7	1.2	923	1	STA2_MOUSE
88	7	1.2	958	1	TRH_DROME
89	7	1.2	1064	1	SYIC_SCHPO
90	7	1.2	1364	1	CND1_XENLA
91	7	1.2	1648	1	YJ9H_YEAST
92	7	1.2	2029	1	LAR_DROME
93	7	1.2	2158	1	MY9B_HUMAN
94	7	1.2	2485	1	POLN_BEVU3
95	7	1.2	2492	1	POLN_BEVVP
96	7	1.2	2492	1	POLN_BEVVT
97	7	1.2	3010	1	POLG_HCVBK
98	7	1.2	3010	1	POLG_HCVJA
99	7	1.2	3010	1	POLG_HCVJT
100	7	1.2	3011	1	POLG_HCV1

ALIGNMENTS

RESULT 1

ALBU HUMAN
ID ALBU HUMAN STANDARD; PRT; 609 AA.
AC P02768; O95574; Q13140; Q9P157; Q9P117; Q9UHS3; Q9UZO3;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86196112; PubMed=3009475;
RA Minigetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,
RA Beattie W.G., Dugaiczky A.;
RT "Molecular structure of the human albumin gene is revealed by
RT nucleotide sequence within q11-22 of chromosome 4.";
RL J. Biol. Chem. 261:6747-6757(1986).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT LYS-420.
RX MEDLINE=82081982; PubMed=6171778;
RA Lawn R.M., Adelman J., Sock S.C., Franke A.E., Houck C.M.,
RA Najarian R.C., Seeburg P.H., Wion K.L.;
RT "The sequence of human serum albumin cDNA and its expression in E.
RT coli.";
RL Nucleic Acids Res. 9:6103-6114(1981).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANT GLY-121.
RX MEDLINE=82105994; PubMed=6275391;
RA Dugaiczky A., Law S.W., Dennison O.E.;
RT "Nucleotide sequence and the encoded amino acids of human serum
RT albumin mRNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).
RC TISSUE=Petal liver;
RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,
RA Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;
RT "Functional prediction of the coding sequences of 121 new genes
RT deduced by analysis of cDNA clones from human fetal liver.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.
RA Huang M.C., Wu H.T.;
RT "The cDNA sequences of human serum albumin.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver, and Skeletal muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Srausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Carninci P., Prange C.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Aravind P., Mullany S.J.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Besak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Munzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Heiton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [8]
RP SEQUENCE OF 25-609.
RX MEDLINE=76187907; PubMed=1225573;
RA Meloun B., Moravsek L., Kostka V.;
RT "Complete amino acid sequence of human serum albumin.";
RL FEBS Lett. 58:134-137(1975).
RN [9]
RP SEQUENCE OF 25-609.
RA Brown J.R., Shockley P., Behrens P.Q.;
RL (In) Bing D.H. (eds.);
RL The chemistry and physiology of the human plasma proteins, pp.23-40,
RL Pergamon Press, New York (1979).
RN [10]
RP SEQUENCE OF 1-455 FROM N.A.
RC TISSUE=Liver;
RA Menaya J., Farrilla R., Ayuso M.S.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [11]
RP SEQUENCE OF 1-26 FROM N.A.
RX MEDLINE=86140099; PubMed=2419329;
RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;
RT "The human albumin gene. Characterization of the 5' and 3' flanking
RT regions and the polymorphic gene transcripts.";
RL J. Biol. Chem. 261:3244-3251(1986).
RN [12]
RP SEQUENCE OF 222-229.
RX MEDLINE=76257808; PubMed=955075;
RA Walker J.E.;
RT "Lysine residue 199 of human serum albumin is modified by
RT acetylsalicylic acid.";
RL FEBS Lett. 66:173-175(1976).
RN [13]
RP SEQUENCE OF 25-44 AND 480-499.
RC TISSUE=Heart;
RX MEDLINE=95203287; PubMed=7895732;
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
RT "The human myocardial two-dimensional gel protein database: update
RT 1994.";
RL Electrophoresis 15:1459-1465(1994).
RN [14]
RP DISULFIDE BONDS.
RA Saber M.A., Stockbauer P., Moravsek L., Meloun B.;
RT "Disulfide bonds in human serum albumin.";
RL Collect. Czech. Chem. Commun. 42:564-579(1977).
RN [15]
RP BILIRUBIN-BINDING SITE.
RX MEDLINE=78186630; PubMed=656055;
RA Jacobsen C.;
RT "Lysine residue 240 of human serum albumin is involved in high-
RT affinity binding of bilirubin.";
RL Biochem. J. 171:453-459(1978).
RN [16]
RP VARIANT CANTERBURY ASN-337.
RX MEDLINE=87157744; PubMed=3828358;
RA Brennan S.O., Herbert P.;
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second
RT domain of serum albumin.";
RL Biochim. Biophys. Acta 912:191-197(1987).
RN [17]
RP VARIANTS NAG-2 AND NAG-3.
RX MEDLINE=88068523; PubMed=3479777;
RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,
RA Satoh C., Neel J.V.;
RT "Amino acid substitutions in inherited albumin variants from
RT Amerindian and Japanese populations.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).
RN [18]
RP VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.
RX MEDLINE=89345611; PubMed=2762316;

RA Arai K., Madison J., Huss K., Ishio K., Satoh C., Fujita M.,
RA Neel J.V., Sakurabayashi I., Putnam F.W.;
RA "Point substitutions in Japanese allolbumins";
RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).
[19]
RP VARIANTS MANAUS; OSAKA; NAGAYA; FUKUOKA; HONOLULU AND NEW-GUINEA.
RX MEDLINE=90115905; PubMed=2404284;
RA Arai K., Madison J., Shimizu A., Putnam F.W.;
RA "Point substitutions in albumin genetic variants from Asia";
RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).
[20]
RN DESCRIPTION OF VARIANT REDHILL.
RX MEDLINE=90115852; PubMed=2104980;
RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;
RA "Albumin Redhill (-1 Arg, 320 Ala--Thr): a glycoprotein variant of
RT human serum albumin whose precursor has an aberrant signal peptidase
RT cleavage site.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).
[21]
RN VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.
RX MEDLINE=91062352; PubMed=2247440;
RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,
RA Watkins S., Putnam F.W.;
RT "Mutations in genetic variants of human serum albumin found in
RT Italy.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).
[22]
RN VARIANT VENEZIA.
RX MEDLINE=91296740; PubMed=2068071;
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,
RA Minchiotti L., Putnam F.W.;
RT "A donor splice mutation and a single-base deletion produce two
RT carboxyl-terminal variants of human serum albumin.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).
[23]
RN VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;
RP KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.
RX MEDLINE=92052189; PubMed=1946412;
RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,
RA Matsuda Y.-I., Anaki I., Putnam F.W.;
RT "Genetic variants of serum albumin in Americans and Japanese.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).
[24]
RN VARIANT CASEBROOK ASN-518.
RX MEDLINE=91316157; PubMed=1859851;
RA Peach R.J., Brennan S.O.;
RT "Structural characterization of a glycoprotein variant of human serum
RT albumin; albumin Casebrook (494 Asp-->Asn).";
RL Biochim. Biophys. Acta 1097:49-54(1991).
[25]
RN VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.
RX MEDLINE=92190239; PubMed=1347703;
RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,
RA Rochu D., Porta F.;
RT "Two allolbumins with identical electrophoretic mobility are produced

Query Match 100.0%; Score 585; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAKSEVAHFRDGLGEENFKALVLIATFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
DB 25 DAKSEVAHFRDGLGEENFKALVLIATFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 84
QY 61 NDKSLHTLFGDKLTVAITRLTYGEMADCCAKQEPNECFLOHKDNPMLPRLVPEV 120
DB 85 NDKSLHTLFGDKLTVAITRLTYGEMADCCAKQEPNECFLOHKDNPMLPRLVPEV 144
QY 121 DVNCTAFHNEETFLKKLYEATRHPPFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
DB 145 DVNCTAFHNEETFLKKLYEATRHPPFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 204
QY 181 KLDELDEGKASSAKORLKCASLQKFGERAFAKAWARLSORFFKAEFVSKLVTLTK 240

DB 205 KLDELDEGKASSAKORLKCASLQKFGERAFAKAWARLSORFFKAEFVSKLVTLTK 264
QY 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDMPA 300
DB 265 VHTTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDMPA 324
QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMLFYEYARRHPDYSVLLLLRLAKTYETTLK 360
DB 325 DLPSLAADFVESKDVCKNVAEAKDVLGMLFYEYARRHPDYSVLLLLRLAKTYETTLK 384
QY 361 CAAADPHECVAKVDFBPKLVPEPQNLKQNCLEPQLGEYKFNALLVRYTKKVPQVST 420
DB 385 CAAADPHECVAKVDFBPKLVPEPQNLKQNCLEPQLGEYKFNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKHPEAKMPCAEADYLSVNLQCLVLEHKTVPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGKVGSKCKHPEAKMPCAEADYLSVNLQCLVLEHKTVPVSDRVTKCCTES 504
QY 481 LVNRFCFSALEVDITYVPKEPNAETFTFHADICTLSKERQIKQTALVELVKGKPKAT 540
DB 505 LVNRFCFSALEVDITYVPKEPNAETFTFHADICTLSKERQIKQTALVELVKGKPKAT 564
QY 541 KEQLKAVMDDPAAAFVCKCKCKADDKTCFAEKGKLVAAASQAALGL 585
DB 565 KEQLKAVMDDPAAAFVCKCKCKADDKTCFAEKGKLVAAASQAALGL 609

RESULT 2

ALBU MACMU
ID _ALBU MACMU STANDARD; PRT; 600 AA.
AC Q28522;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor (Fragment).
GN ALB
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91211971; PubMed=8460152;
RA Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,
RA Dwulet J., Putnam F.W.;
RT "cDNA and protein sequence of polymorphic macaque albumins that differ
RT in bilirubin binding.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).
CC !- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC !- SUBCELLULAR LOCATION: Secreted.
CC !- TISSUE SPECIFICITY: Plasma.
CC !- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC !- SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M90463; AAA36906.1; -
DR PIR; A47391; A47391.
DR HSSP; P02768; 1E7B.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.

ProDom; PD002486; Serum albumin; 1.
DR SMART; SMO0103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT NON TER 1
FT SIGNAL 1 10 BY SIMILARITY.
FT PROPEP 11 16 BY SIMILARITY.
FT CHAIN 17 600 SERUM ALBUMIN.
FT DOMAIN 17 137 ALBUMIN 1.
FT DOMAIN 204 389 ALBUMIN 2.
FT DOMAIN 396 587 ALBUMIN 3.
FT METAL 19 19 COPPER (BY SIMILARITY).
FT BINDING 256 256 BILIRUBIN (POTENTIAL).
FT DISULFID 69 78 BY SIMILARITY.
FT DISULFID 91 107 BY SIMILARITY.
FT DISULFID 106 117 BY SIMILARITY.
FT DISULFID 140 185 BY SIMILARITY.
FT DISULFID 184 193 BY SIMILARITY.
FT DISULFID 216 262 BY SIMILARITY.
FT DISULFID 261 269 BY SIMILARITY.
FT DISULFID 281 295 BY SIMILARITY.
FT DISULFID 294 305 BY SIMILARITY.
FT DISULFID 332 377 BY SIMILARITY.
FT DISULFID 376 385 BY SIMILARITY.
FT DISULFID 408 454 BY SIMILARITY.
FT DISULFID 453 464 BY SIMILARITY.
FT DISULFID 477 493 BY SIMILARITY.
FT DISULFID 492 503 BY SIMILARITY.
FT DISULFID 530 575 BY SIMILARITY.
FT DISULFID 574 583 BY SIMILARITY.
SQ SEQUENCE 600 AA; 67880 MW; E45C871A670E740B CRC64;

Query Match 12.8%; Score 75; DB 1; Length 600;
Best Local Similarity 100.0%; Pred. No. 3.7e-68;
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 HVKLNVETFEAKTVADESSENCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPR 98
DB 55 HVKLNVETFEAKTVADESSENCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPR 114

QY 99 NECFLQKODNPNLP 113
DB 115 NECFLQKODNPNLP 129

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

DR EMBL; X84842; CAA59279.1; -.
DR PIR; JC4660; S57632.
DR HSP; P02768; 1E7B.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PRO0802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SMO0103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
FT SIGNAL 1 18 BY SIMILARITY.
FT PROPEP 19 24 BY SIMILARITY.
FT CHAIN 25 608 SERUM ALBUMIN.
FT DOMAIN 25 205 ALBUMIN 1.
FT DOMAIN 212 397 ALBUMIN 2.
FT DOMAIN 404 595 ALBUMIN 3.
FT METAL 27 27 COPPER.
FT DISULFID 77 86 BY SIMILARITY.
FT DISULFID 99 115 BY SIMILARITY.
FT DISULFID 114 125 BY SIMILARITY.
FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 192 201 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 340 385 BY SIMILARITY.
FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 485 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 538 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
SQ SEQUENCE 608 AA; 68659 MW; 07E629CAC5P60E5F CRC64;

Query Match 5.0%; Score 29; DB 1; Length 608;
Best Local Similarity 100.0%; Pred. No. 3.6e-21;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 ECCHGDLLECCADDRADLAKYICENQDSIS 272
DB 268 ECCHGDLLECCADDRADLAKYICENQDSIS 296

RESULT 4
ALBU_HORSE
ID ALBU_HORSE STANDARD; PRT; 607 AA.
AC P35747;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Equ c 3).
GN ALB.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Ferissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RC TISSUE=Liver;
RX MEDLINE=93345495; PubMed=8344282;
RA Ho J.X., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;
RT "X-ray and primary structure of horse serum albumin (Equus caballus)

at 0.27-nm resolution.";
Eur. J. Biochem. 215:205-212(1993).
-!- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.
-!- SUBCELLULAR LOCATION: Secreted.
-!- TISSUE SPECIFICITY: Plasma.
-!- ALLERGEN: Causes an allergic reaction in human. Binds IgG.
-!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
-!- SIMILARITY: Contains 3 albumin domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X74045; CAA52194.1; -;
PIR; S34053; ABOH.
HSP; P02768; 1E7B.
InterPro; IPR000264; Serum albumin.
Pfam; PF00273; transport prot; 3.
PRINTS; PR00802; SERUMALBUMIN.
ProDom; PD002486; Serum_albumin; 1.
SMART; SM00103; ALBUMIN; 3.
PROSITE; PS00212; ALBUMIN; 3.
Metal-binding; Lipid-binding; Repeat; Signal; Allergen.
SIGNAL 1 18 BY SIMILARITY.
DR PROPEP 19 24 SERUM ALBUMIN.
FT CHAIN 25 607 BY SIMILARITY.
FT DOMAIN 25 204 ALBUMIN 1.
FT DOMAIN 211 396 ALBUMIN 2.
FT DOMAIN 403 594 ALBUMIN 3.
FT METAL 27 27 COPPER (BY SIMILARITY).
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 147 192
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 288 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461
FT DISULFID 460 471
FT DISULFID 484 500
FT DISULFID 499 510
FT DISULFID 537 582
FT DISULFID 581 590
SQ SEQUENCE 607 AA; 68598 MW; 256F6E830A1B90C5 CRC64;

Query Match 4.6%; Score 27; DB 1; Length 607;
Best Local Similarity 100.0%; Pred.No. 4e-19;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 ADESAEKCCKSLHTFGDKLCTVATLR 81
Db 79 ADESAEKCCKSLHTFGDKLCTVATLR 105

RESULT 5
ALBU PIG STANDARD; PRT; 605 AA.
AC P08335; Q28018;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor (Fragment).

GN ALB.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
CX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA MEDLINE=89016582; PubMed=3174440;
RA Baldwin G.S., Weinstein J.;
RT "Nucleotide sequence of porcine liver albumin.";
RL Nucleic Acids Res. 16:9045-9045(1988).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X12422; CAA30970.1; -;
EMBL; M36787; AAA30988.1; -;
PIR; S01382; ASFGS.
HSP; P02768; 1E7H.
InterPro; IPR000264; Serum albumin.
Pfam; PF00273; transport prot; 3.
PRINTS; PR00802; SERUMALBUMIN.
ProDom; PD002486; Serum_albumin; 1.
SMART; SM00103; ALBUMIN; 3.
PROSITE; PS00212; ALBUMIN; 3.
Metal-binding; Lipid-binding; Repeat; Signal; Copper.
NON TER 1 1
SIGNAL <1 16 BY SIMILARITY.
FT PROPEP 17 22 BY SIMILARITY.
FT CHAIN 23 605 SERUM ALBUMIN.
FT DOMAIN 23 202 ALBUMIN 1.
FT DOMAIN 209 394 ALBUMIN 2.
FT DOMAIN 401 592 ALBUMIN 3.
FT METAL 31 31 COPPER (BY SIMILARITY).
FT DISULFID 75 84 BY SIMILARITY.
FT DISULFID 97 113 BY SIMILARITY.
FT DISULFID 112 123 BY SIMILARITY.
FT DISULFID 145 190 BY SIMILARITY.
FT DISULFID 189 198 BY SIMILARITY.
FT DISULFID 221 267 BY SIMILARITY.
FT DISULFID 266 274 BY SIMILARITY.
FT DISULFID 286 300 BY SIMILARITY.
FT DISULFID 299 310 BY SIMILARITY.
FT DISULFID 337 382 BY SIMILARITY.
FT DISULFID 381 390 BY SIMILARITY.
FT DISULFID 413 459 BY SIMILARITY.
FT DISULFID 458 469 BY SIMILARITY.
FT DISULFID 482 498 BY SIMILARITY.
FT DISULFID 497 508 BY SIMILARITY.
FT DISULFID 535 580 BY SIMILARITY.
FT DISULFID 579 588 BY SIMILARITY.
FT CONFLICT 562 562 E -> D (IN REF. 1; AAA30988).
SQ SEQUENCE 605 AA; 69410 MW; 3E556BDD1A1F4FF CRC64;

Query Match 4.4%; Score 26; DB 1; Length 605;
Best Local Similarity 100.0%; Pred.No. 4.2e-18;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 ECCHGDLLCADDRAADLAKYICENQD 269

|||||
 265 ECCHGDLLECAADRADLAKYICENQD 290

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DB ID ALBU RAT STANDARD; PRT; 608 AA.

AC P02770; P11382;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor [Contains: Neurotensin-related peptide (NRP)].
 GN ALB.

OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. PubMed=7017712;
 RX MEDLINE=81223722; PubMed=7017712;
 RA Sargent T.D., Yang M., Bonner J.;
 RT "Nucleotide sequence of cloned rat serum albumin messenger RNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 78:243-246 (1981).
 RN [2]
 RP SEQUENCE OF 1-38, AND PROCESSING.
 RX MEDLINE=77249657; PubMed=893447;
 RA Strauss A.W., Bennett C.D., Donohue A.M., Rodkey J.A., Alberts A.W.;
 RT "Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analysis of the direct translation product of albumin messenger RNA.";
 RL J. Biol. Chem. 252:6846-6855 (1977).
 RN [3]
 RP SEQUENCE OF 25-222.
 RX MEDLINE=78109429; PubMed=564345;
 RA Isemura S., Ikenaka T.;
 RT "Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavage of rat serum albumin.";
 RL J. Biochem. 83:35-48 (1978).
 RN [4]
 RP SEQUENCE OF 223-288 AND 572-608.
 RX MEDLINE=76260153; PubMed=956149;
 RA Isemura S., Ikenaka T.;
 RT "Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino acid sequences of four fragments.";
 RL J. Biochem. 79:1183-1196 (1976).
 RN [5]
 RP SEQUENCE OF 166-174.
 RX TISSUE=Plasma;
 RC MEDLINE=87194605; PubMed=2437111;
 RA Cartaway R.E., Mitra S.P., Cochrane D.E.;
 RT "Structure of a biologically active neurotensin-related peptide obtained from pepsin-treated albumin(s).";
 RL J. Biol. Chem. 262:5968-5973 (1987).
 RN [6]
 RP COPPER-BINDING
 RX MEDLINE=79001617; PubMed=80265;
 RA Aoyagi Y., Ikenaka T., Ichida F.;
 RT "Copper(II)-binding ability of human alpha-fetoprotein.";
 RC Cancer Res. 38:3483-3486 (1978).
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.
 CC -1- FUNCTION: NRP regulates fat digestion, lipid absorption, and blood flow (Potential).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Plasma.
 CC -1- SIMILARITY: Belongs to the ALB/AFP/VDB family.
 CC -1- SIMILARITY: Contains 3 albumin domains.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

EMBL: V01222; CAA24532.1; --
 DR PIR: A93872; ABETS.
 DR HSP: P02768; IE7E.
 DR InterPro: IPR000284; Serum_albumin.
 DR Pfam: PF00273; transport_prot; 3.
 DR PRINTS: PR00802; SERUMALBUMIN.
 DR ProDom: PD002486; Serum_albumin; 1.
 DR SMART: SM00103; ALBUMIN; 3.
 DR ProSITE: PS00212; ALBUMIN; 3.
 DR Metal-binding; Lipid-binding; Repeat; Signal; Copper.
 KW METAL-BINDING; LIPID-BINDING; REPEAT; SIGNAL; COPPER.
 FT SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 608
 FT PEPTIDE 166 174 SERUM ALBUMIN.
 FT DOMAIN 25 205 NEUTROSENSIN-RELATED PEPTIDE.
 FT DOMAIN 212 397 ALBUMIN 1.
 FT DOMAIN 404 595 ALBUMIN 2.
 FT METAL 27 27 ALBUMIN 3.
 FT DISULFID 77 86 COPPER.
 FT DISULFID 99 115 BY SIMILARITY.
 FT DISULFID 114 125 BY SIMILARITY.
 FT DISULFID 148 193 BY SIMILARITY.
 FT DISULFID 192 201 BY SIMILARITY.
 FT DISULFID 224 270 BY SIMILARITY.
 FT DISULFID 269 277 BY SIMILARITY.
 FT DISULFID 289 303 BY SIMILARITY.
 FT DISULFID 302 313 BY SIMILARITY.
 FT DISULFID 340 385 BY SIMILARITY.
 FT DISULFID 384 393 BY SIMILARITY.
 FT DISULFID 416 462 BY SIMILARITY.
 FT DISULFID 461 472 BY SIMILARITY.
 FT DISULFID 485 501 BY SIMILARITY.
 FT DISULFID 500 511 BY SIMILARITY.
 FT DISULFID 538 583 BY SIMILARITY.
 FT DISULFID 582 591 BY SIMILARITY.
 FT VARIANT 262 262 V -> L.
 FT CONFLICT 174 174 Y -> L (IN REF. 5).
 SQ SEQUENCE 608 AA; 68718 MW; 5BB497A282411AB7 CRC64;

Query Match 4.4%; Score 26; DB 1; Length 608;
 Best Local Similarity 100.0%; Pred. No. 4,2e-18;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 88 ADCCAKQEPERNECFLOHKDDNPNLP 113
 DB 112 ADCCAKQEPERNECFLOHKDDNPNLP 137

RESULT 7

ALBU MERUN STANDARD; PRT; 609 AA.
 ID ALBU MERUN
 AC O35090;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor.
 GN ALB.
 OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
 OC Meriones
 OX NCBI_TaxID=10047;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MGS IDR; TISSUE=Liver;
 RX MEDLINE=98116663; PubMed=9455485;
 RA Yoshida K., Sato-Oshima A., Sinozawa H.;
 RT "Sequencing of cDNA encoding serum albumin and its extrahepatic

RT synthesis in the Mongolian gerbil, Meriones unguiculatus.";

RL DNA Res. 4:351-354(1997).

CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good

CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,

CC hormones, bilirubin and drugs. Its main function is the regulation

CC of the colloidal osmotic pressure of blood.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Plasma.

CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.

CC -!- SIMILARITY: Contains 3 albumin domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AB006197; BAA21765.1; -.

DR PIR; JCS838; JCS838.

DR HSSP; P02768; 1E7B.

DR InterPro; IPR000264; Serum_albumin.

DR Pfam; PF00273; transport_prot; 3.

DR PRINTS; PR00802; SERUMALBUMIN.

DR ProDom; PD002486; Serum_albumin; 1.

DR SMART; SM00103; ALBUMIN; 3.

DR PROSITE; PS00212; ALBUMIN; 3.

DR Metal-binding; Lipid-binding; Repeat; Signal; Copper.

FT SIGNAL 1 18

FT PROPEP 19 24

FT CHAIN 25 609

FT DOMAIN 25 206

FT DOMAIN 213 398

FT DOMAIN 405 596

FT METAL 28 28

FT DISULFID 78 87

FT DISULFID 100 116

FT DISULFID 115 126

FT DISULFID 149 194

FT DISULFID 193 202

FT DISULFID 225 271

FT DISULFID 270 278

FT DISULFID 290 304

FT DISULFID 303 314

FT DISULFID 341 386

FT DISULFID 385 394

FT DISULFID 417 463

FT DISULFID 462 473

FT DISULFID 486 502

FT DISULFID 501 512

FT DISULFID 539 584

FT DISULFID 583 592

FT DISULFID 609 AA; 68940 MW; 9CA5F9767EF1A48 CRC64;

Query Match 4.3%; Score 25; DB 1; Length 609;

Best Local Similarity 100.0%; Pred. No. 4.4e-17;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 EVADCCAKOEPRNECFQHKDDNP 110

DB 111 EVADCCAKOEPRNECFQHKDDNP 135

RESULT 8

ALBU_RABIT STANDARD; PRT; 608 AA.

AC P49065;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Serum albumin precursor.

GN ALB.

OS Oryctolagus cuniculus (Rabbit).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

CC NCBI_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=New Zealand white; TISSUE=Liver;

RA Sheffield W.P., Syed S., Schuyler P.D.;

RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good

CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,

CC hormones, bilirubin and drugs. Its main function is the regulation

CC of the colloidal osmotic pressure of blood.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Plasma.

CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.

CC -!- SIMILARITY: Contains 3 albumin domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U18344; AAB58347.1; -.

DR HSSP; P02768; 1E7B.

DR InterPro; IPR000264; Serum_albumin.

DR Pfam; PF00273; transport_prot; 3.

DR PRINTS; PR00802; SERUMALBUMIN.

DR ProDom; PD002486; Serum_albumin; 1.

DR SMART; SM00103; ALBUMIN; 3.

DR PROSITE; PS00212; ALBUMIN; 3.

DR Metal-binding; Lipid-binding; Repeat; Signal; Copper.

FT SIGNAL 1 18

FT PROPEP 19 24

FT CHAIN 25 608

FT DOMAIN 25 205

FT DOMAIN 212 397

FT DOMAIN 404 595

FT METAL 27 27

FT DISULFID 77 86

FT DISULFID 99 115

FT DISULFID 114 125

FT DISULFID 148 193

FT DISULFID 192 201

FT DISULFID 224 270

FT DISULFID 269 277

FT DISULFID 289 303

FT DISULFID 302 313

FT DISULFID 340 385

FT DISULFID 384 393

FT DISULFID 416 462

FT DISULFID 461 472

FT DISULFID 485 501

FT DISULFID 500 511

FT DISULFID 538 583

FT DISULFID 582 591

FT DISULFID 608 AA; 68914 MW; CF5E92647AAFE9A2 CRC64;

Query Match 3.9%; Score 23; DB 1; Length 608;

Best Local Similarity 100.0%; Pred. No. 4.9e-15;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 403 FQNALVRYTKKVPQVSTPTIVE 425

DB 427 FQNALVRYTKKVPQVSTPTIVE 449

RESULT 9

ALBU_BOVIN STANDARD; PRT; 607 AA.

ID ALBU_BOVIN

AC P02769; 002787;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Bos d 6).
GN ALB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
CX NCBI_TaxID=9913;
RN SEQUENCE FROM N.A.
RP Holowachuk E.W., Stoltenberg J.K., Reed R.G., Peters T. Jr.;
RA Submitted (AUG-1991) to the EMBL/GenBank/DBJ databases.
RL [2]
RN SEQUENCE FROM N.A., AND VARIANT THR-214.
RP TISSUE=Liver;
RC Barry T., Power S., Gannon F.;
RA Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
RL [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Liver;
RC Hilger C., Grigioni F., de Beaufort C., Michel G., Hentges F.;
RA Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RL [4]
RN SEQUENCE FROM N.A., AND VARIANT THR-214.
RP Wu H.T., Huang M.C.;
RA "The complete cDNA sequence of bovine serum albumin.";
RT Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RL [5]
RN SEQUENCE OF 1-32.
RX MEDLINE=80024278; PubMed=488109;
RA McGilivray R.T.A., Chung D.W., Davie E.W.;
RT "Biosynthesis of bovine plasma proteins in a cell-free system. Amino-terminal sequence of preproalbumin.";
RL Eur. J. Biochem. 98:477-485 (1979).
RN [6]
RN SEQUENCE OF 25-424 AND 429-607, AND VARIANT THR-214.
RA Brown J.R.;
RT "Structure of bovine serum albumin.";
RL Fed. Proc. 34:591-591 (1975).
RN [7]
RN REVISIONS TO 190-195.
RA Brown J.R.;
RL Submitted (APR-1975) to the PIR data bank.
RN [8]
RN SEQUENCE OF 402-433.
RX MEDLINE=8203364; PubMed=7283978;
RA Reed R.G., Putnam F.W., Peters T. Jr.;
RT "Sequence of residues 400-403 of bovine serum albumin.";
RL Biochem. J. 191:867-868 (1980).
RN [9]
RN SEQUENCE OF 19-28.
RX MEDLINE=77134075; PubMed=843354;
RA Patterson J.E., Geller D.M.;
RT "Bovine microsomal albumin: amino terminal sequence of bovine proalbumin.";
RL Biochem. Biophys. Res. Commun. 74:1220-1226 (1977).
RN [10]
RN SEQUENCE, AND REVISIONS TO 118-119 AND 180.
RX MEDLINE=91083649; PubMed=2260975;
RA Hirayama K., Akashi S., Furuya M., Fukuhara K.-I.;
RT "Rapid confirmation and revision of the primary structure of bovine serum albumin by ESI/MS and Frit-FAB LC/MS.";
RL Biochem. Biophys. Res. Commun. 173:639-646 (1990).
RN [11]
RN SEQUENCE OF 25-41.
RX MEDLINE=88267456; PubMed=3389500;
RA Hsieh J.C., Lin F.P., Tam M.P.;
RT "Electroblotting onto glass-fiber filter from an analytical isoelectrofocusing gel: a preparative method for isolating proteins for N-terminal microsequencing.";
RT

RL Anal. Biochem. 170:1-8 (1988).
RN [12]
RP SEQUENCE OF 437-451.
RA Vilbois F.;
RL Submitted (AUG-1998) to Swiss-Prot.
RN [13]
RP DISULFIDE BONDS.
RA Brown J.R.;
RT "Structure of serum albumin: disulfide bridges.";
RL Fed. Proc. 33:1389-1389 (1974).
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M73993; AAA51411.1; -;
CC EMBL; X53989; CAA41735.1; -;
CC EMBL; I17769; CAA76847.1; -;
CC EMBL; AF542068; AAN17824.1; -;
CC HSSP; P02768; 1E7B.
CC Interpro; IPR000264; Serum_albumin.
CC Pfam; PF00273; transport_prot; 3.
CC PRINTS; PR00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum_albumin; 1.
CC SMART; SM00103; ALBUMIN_3.
CC PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen;
FT Polymorphism.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 607 SERUM ALBUMIN.
FT DOMAIN 25 204 ALBUMIN 1.
FT DOMAIN 211 396 ALBUMIN 2.
FT DOMAIN 403 594 ALBUMIN 3.
FT METAL 27 27 COPPER (BY SIMILARITY).
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 147 192
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 288 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461
FT DISULFID 460 471
FT DISULFID 484 500
FT DISULFID 499 510
FT DISULFID 537 582
FT DISULFID 581 590
FT VARIANT 214 214
FT CONFLICT 302 302 C -> T.
FT CONFLICT 304 305 KP -> PC (IN REF. 6).
FT CONFLICT 324 324 N -> D (IN REF. 6).
FT CONFLICT 394 395 ST -> TS (IN REF. 6).
FT CONFLICT 437 437 K -> R (IN REF. 12).
FT CONFLICT 493 494 SE -> ES (IN REF. 6).
FT CONFLICT 493 494
FT SEQUENCE 607 AA; 69293 MW; 39167DFE768585D4 CRC64;

```
Query Match          3.8%; Score 22; DB 1; Length 607;
Best Local Similarity 100.0%; Pred. No. 5.1e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 244 ECCHGDLLECCADRADLAKYIC 265
Db 267 ECCHGDLLECCADRADLAKYIC 288

RESULT 10
ALBU SHEEP
ID ALBU SHEEP STANDARD; PRT; 607 AA.
AC P14639;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.

OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=9009888; PubMed=2602160;
RA Brown W.M., Dziegielewska K.M., Foreman R.C., Saunders N.R.;
RT "Nucleotide and deduced amino acid sequence of sheep serum albumin.";
RL Nucleic Acids Res. 17:10495-10495(1989).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X17055; CAA34903.1; -.
CC PIR; S06936; ABSHS.
CC HSSP; P02768; 1E7B.
CC InterPro; IPR000264; Serum albumin.
CC Pfam; PF00273; transport prot.; 3.
CC PRINTS; PR00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum albumin; 1.
CC SMART; SM00103; ALBUMIN; 3.
CC KX Metal-binding; Lipid-binding; Repeat; Signal; Copper.
CC PROSITE; PS00212; ALBUMIN; 3.
CC SIGNAL 1 18
CC FT PROPEP 19 24 BY SIMILARITY.
CC FT CHAIN 25 607 SERUM ALBUMIN.
CC FT DOMAIN 25 204 ALBUMIN 1.
CC FT DOMAIN 211 396 ALBUMIN 2.
CC FT DOMAIN 403 594 ALBUMIN 3.
CC FT METAL 27 27 COPPER (BY SIMILARITY).
CC FT DISULFID 77 86 BY SIMILARITY.
CC FT DISULFID 99 115 BY SIMILARITY.
CC FT DISULFID 114 126 BY SIMILARITY.
CC FT DISULFID 147 152 BY SIMILARITY.
CC FT DISULFID 191 200 BY SIMILARITY.
CC FT DISULFID 223 269 BY SIMILARITY.
CC FT DISULFID 268 276 BY SIMILARITY.
CC FT DISULFID 288 302 BY SIMILARITY.
```

```
FT DISULFID 301 312 BY SIMILARITY.
FT DISULFID 339 384 BY SIMILARITY.
FT DISULFID 383 392 BY SIMILARITY.
FT DISULFID 415 461 BY SIMILARITY.
FT DISULFID 460 471 BY SIMILARITY.
FT DISULFID 484 500 BY SIMILARITY.
FT DISULFID 499 510 BY SIMILARITY.
FT DISULFID 537 582 BY SIMILARITY.
FT DISULFID 581 590 BY SIMILARITY.
SQ SEQUENCE 607 AA; 69188 MW; 84979A87F8B86596 CRC64;

Query Match          3.8%; Score 22; DB 1; Length 607;
Best Local Similarity 100.0%; Pred. No. 5.1e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 244 ECCHGDLLECCADRADLAKYIC 265
Db 267 ECCHGDLLECCADRADLAKYIC 288

RESULT 11
ALBU CANFA
ID ALBU CANFA STANDARD; PRT; 608 AA.
AC P49822; O77705; Q9TS24;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Can f 3).
GN ALB.

OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Hilger C.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=20148667; PubMed=10669848;
RA Pandjaitan B., Swoboda I., Brandejsky-Pichler F., Rumpold H.,
RA Valenta R., Spitzauer S.;
RT "Escherichia coli expression and purification of recombinant dog
RT albumin, a cross-reactive animal allergen.";
RL J. Allergy Clin. Immunol. 105:279-285(2000).
RN [3]
RP SEQUENCE OF 25-48.
RX MEDLINE=75011422; PubMed=4414013;
RA Dixon J.W., Sarkar B.;
RT "Isolation, amino acid sequence and copper(II)-binding properties of
RT peptide (1-24) of dog serum albumin.";
RN [4]
RP SEQUENCE OF 25-38.
RC TISSUE=Heart;
RX MEDLINE=98163340; PubMed=9504812;
RA Dunn M.J., Corbett J.M., Wheeler C.H.;
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT dog heart proteins.";
RL Electrophoresis 18:2795-2802(1997).
RN [5]
RP SEQUENCE OF 215-478 FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=94201492; PubMed=7512102;
RA Spitzauer S., Schweiger C., Sperr W.R., Pandjaitan B., Valent P.,
RA Muehl S., Ebner C., Scheiner O., Kraft D., Rumpold H.;
RT "Molecular characterization of dog albumin as a cross-reactive
RT allergen.";
RL J. Allergy Clin. Immunol. 93:614-627(1994).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
```

D	ALBU MOUSE	STANDARD;	PRT;	608 AA.
C	207724;	Q61802;		
CT	01-APR-1988	(Rel. 07, Created)		
CT	15-JUL-1999	(Rel. 38, Last sequence update)		
WT	15-MAR-2004	(Rel. 43, Last annotation update)		
DE	Serum albumin precursor.			
DE	ALB OR ALB1 OR ALB-1.			
SS	Mus musculus (Mouse).			
CC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
XX	NCBI_TaxId=10090;			
XX	[1]			
XX	SEQUENCE FROM N.A.			
XX	TISSUE=Liver;			
CC	van Reeth T., Gabant P., Dreze P., Szpirer J., Szpirer C.;			
XX	Submitted (Sep-1998) to the EMBL/GenBank/DBJ databases.			

[2] SEQUENCE FROM N.A.
STRAIN=C57BL/6J; TISSUE=Testis;
MEDLINE=21085660; PubMed=11217851;
Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Konno H., Kondo S., Yamakura S.,
Aizawa M., Nishi K., Kiyosawa H., Kadoya T., Bono H., Kasukawa T., Saito R.,
Saito T., Okazaki Y., Gojobori T., Bono H., Matsuda H., Ashburner M., Batalov S., Casavant T.,
Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
Kuehl P., Lewis S., Masuzo Y., Nikaide I., Pesole G., Quackenbush J.,
Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
Akai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustincich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kontuski S.,
Hayashizaki Y.,
"Functional annotation of a full-length mouse cDNA collection."
Nature 409:685-690(2001).
[3] SEQUENCE FROM N.A.
STRAIN=FVB/N; TISSUE=Liver;
MEDLINE=22398257; PubMed=12477932;
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Moore T., Max S.I., Wang J., Hsieh F.,
Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
Boasak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smalios D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
"Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences."
Proc Natl Acad Sci U.S.A. 99:16899-16903(2002).

RT human and mouse cDNA sequences.";
Proc Natl Acad Sci U.S.A. 99:16899-16903 (2002).

RX MEDLINE=88216123; PubMed=2452956;

RT "The rate of molecular evolution of alpha-fetoprotein approaches that

RN [5]
RP SEQUENCE OF 477-551 FROM N.A.

RN [5]
RP SEQUENCE OF 477-551 FROM N.A.

RA Grossman A., Nicholl J., Antonelli J.,
RA Sutherland G.R., Mak T.W.;
RL Submitted (OCT-1996) to the EMEC

RA Grossman A., Nicholl J., Antonelli
RA Sutherland G.R., Mak T.W.;
RL Submitted (OCT-1996) to the EMEP

RA Grossman A., Nicholl J., Antoni
RA Sutherland G.R., Mak T.W.;
RI Submitted (OCT-1996) to the EME

RP SEQUENCE FROM N.A. (ISOFORMS A; B AND C).
RX MEDLINE=97459673; PubMed=9315633;
RT Zhang L., Pagano J.S.;
RA "IRF-7, a new interferon regulatory factor associated with Epstein-
RT Barr virus latency";
RL Mol. Cell. Biol. 17:5748-5757(1997).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM D).
RX MEDLINE=99003279; PubMed=9786932;
RA Au W.-C., Moore P.A., Lafleur D.W., Tombal B., Pitha P.M.;
RT "Characterization of the interferon regulatory factor-7 and its
RT potential role in the transcription activation of interferon A
RT genes";
RL J. Biol. Chem. 273:29210-29217(1998).
CC -!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR. BINDS TO THE INTERFERON-
CC STIMULATED RESPONSE ELEMENT (ISRE) IN IFN PROMOTERS AND IN THE Q
CC PROMOTER (QP) OF EBV NUCLEAR ANTIGEN-1 (EBNA1).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=A;
CC IsoId=Q92985-1; Sequence=Displayed;
CC Name=B; Synonyms=Beta;
CC IsoId=Q92985-2; Sequence=VSP_002760;
CC Name=C; Synonyms=Gamma;
CC IsoId=Q92985-3; Sequence=VSP_002758, VSP_002759;
CC Name=D; Synonyms=H;
CC IsoId=Q92985-4; Sequence=VSP_002757;
CC -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN SPLEEN, THYMUS, AND
CC PERIPHERAL BLOOD LEUKOCYTES.
CC -!- SIMILARITY: Belongs to the IRF family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U730336; AAB17190.1; -;
DR EMBL; U53830; AAB80686.1; -;
DR EMBL; U53831; AAB80688.1; -;
DR EMBL; U53832; AAB80690.1; -;
DR EMBL; AF076494; AAC70995.1; -;
DR HSSP; P23906; ZIRP.
DR TRANSFAC; T04674; -;
DR TRANSFAC; T05106; -;
DR Genew; HGNC:6122; IRF7.
DR MIM; 605047; -;
DR GO; GO:0005737; Cytoplasm; TAS.
DR GO; GO:0005634; Cytocent; TAS.
DR GO; GO:0003704; P:specific RNA polymerase II transcription fa. . .; TAS.
DR GO; GO:0000122; P:negative regulation of transcription from P. . .; TAS.
DR GO; GO:0009615; P:response to viruses; TAS.
DR InterPro; IPR001346; IRF.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF00605; IRF_1.
DR PRINTS; PR00267; INTERNEGCT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
KW Transcription regulation; DNA-binding; Nuclear protein; Activator;
KW Alternative splicing.
FT DNA_BIND 13 122
FT VARSPLIC 1 6
FT -----
FT TRYPTOPHAN PENTAD REPEAT.
FT MALAPE -> MPVPERPAAGDSRPGR (in isoform
FT D).
FT /FTId=VSP_002757.
FT GPPGPFLLAETHA -> AQGSLGSCGTGGQ (in
FT isoform C).
FT /FTId=VSP_002758.
FT Missing (in isoform C).
FT /FTId=VSP_002759.

FT VARSPLIC 228 256 Missing (in isoform B).
FT /FTId=VSP_002760.
FT E -> K (in REF. 2).
FT CONFLICT 179 179
FT CONFLICT 412 412 Q -> R (in REF. 3).
SQ SEQUENCE 503 AA; 54278 MW; AAG39E0E272727C CRC64;
Query Match 1.4%; Score 8; DB 1; Length 503;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 211 PKAWAVAR 218
DB 60 PKAWAVAR 67
RESULT 17
BUD5_YEAST
ID_BUD5_YEAST STANDARD; PRT; 538 AA.
AC P25300;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Bud site selection protein BUD5.
DE BUD5 OR YCR038C OR YCR38C OR YCR526.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetaceae; Saccharomycetes;
OC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92160397; PubMed=1789011;
RA Jaquet M., Buhler J.-M., Iborra F., Francinques-Gaillard M.-C.,
RA Soustelle C.;
RT "The MAT locus revisited within a 9.8 kb fragment of chromosome III
RT containing BUD5 and two new open reading frames";
RL Yeast 7:881-888(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=SP1;
RX MEDLINE=91292524; PubMed=1905981;
RA Chant J., Corrado K., Pringle J.R., Herskowitz I.;
RT "Yeast BUD5, encoding a putative GDP-GTP exchange factor, is
RT necessary for bud site selection and interacts with bud formation
RT gene BEM1";
RL Cell 65:1213-1224(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=91292525; PubMed=1905982;
RA Powers S., Gonzales E., Christensen T., Cubert J., Broek D.;
RT "Functional cloning of BUD5, a CDC25-related gene from S. cerevisiae
RT that can suppress a dominant-negative RAS2 mutant";
RL Cell 65:1225-1231(1991).
RN [4]
RP SEQUENCE OF 181-538 FROM N.A.
RX STRAIN=S288c / FY1679;
RX MEDLINE=91181345; PubMed=1964349;
RA Thierry A., Fairhead C., Dujon B.;
RT "The complete sequence of the 8.2 kb segment left of MAT on
RT chromosome III reveals five ORFs, including a gene for a yeast
RT ribokinase";
RL Yeast 6:521-534(1990).
CC -!- FUNCTION: Putative GDP-GTP exchange factor. Required to produce
CC both the axial and bipolar patterns of bud site selection. It
CC interacts with bud formation gene BEM1. BUD5 might act on the
CC Ras-like proteins, BUD1/RSR1 and CDC42. It appears to bind to
CC Ras proteins but not to activate the Ras proteins.
CC -!- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.
CC -!- SIMILARITY: Contains 1 Ras-GEF domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M63552; AAA34460.1; -
EMBL; M68938; AAA34462.1; -
EMBL; X59720; CAA42305.1; -
EMBL; X56909; CAA42030.1; -
PIR; S19450; BWBYD5.
GenOnline; 138944; -
SGD; S000634; BUD5.
GO; GO:0005935; C:bud neck; IDA.
GO; GO:0001131; C:incipient bud site; IDA.
InterPro; IPR008937; RasGEF.
InterPro; IPR001895; RasGRF_CDC25.
PIfam; PF00617; RasGEF; 1.
SMART; SM00147; RasGEF; 1.
SMART; SM00229; RasGEF; 1.
PROSITE; PS00720; RasGEF; 1.
PROSITE; PS50009; RasGEF; CAT; 1.
PROSITE; PS50212; RasGEF_YTER; 1.
Guanine-nucleotide releasing factor.
KW DOMAIN 120 235 N-TERMINAL RAS-GEF.
FT DOMAIN 308 536 RAS-GEF.
FT VARIANT 297 297 P -> L.
FT VARIANT 377 377 R -> S.
FT CONFLICT 6 6 R -> P (IN REF. 1).
FT CONFLICT 7 7 MISSING (IN REF. 2).
FT CONFLICT 121 121 V -> D (IN REF. 2).
SQ SEQUENCE 538 AA; 62917 MW; F700F81A78B20DA CRC64;

Query Match 1.4%; Score 8; DB 1; Length 538;
Best Local Similarity 100.0%; Pred. No. 9.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AFAQYLQ 33
DB 485 AFAQYLQ 492

RESULT 18
GGT1_PIG
ID_GGT1_PIG STANDARD; PRT; 568 AA.
AC F20735;
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Gamma-glutamyltranspeptidase 1 precursor (EC 2.3.2.2) (Gamma-glutamyltransferase 1) (GGT 1).
DE GGT1 OR GGT.
GN Sus scrofa (Pig).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain capillary;
RX MEDLINE=89377938; PubMed=2476308;
RA Papadriakopoulou A., Frey A., Gassen H.G.;
RT "Cloning and expression of gamma-glutamyl transpeptidase from isolated porcine brain capillaries."
RL Eur. J. Biochem. 183:693-698(1989).
RN [2]
RP GENE ORGANIZATION AND ALTERNATIVE PROMOTER USAGE.
RX MEDLINE=99320748; PubMed=1032451;
RA Chikhi N., Holic N., Gueliaen G., Laperche Y.;
RT "Gamma-glutamyl transpeptidase gene organization and expression: a comparative analysis in rat, mouse, pig and human species."
RL Comp. Biochem. Physiol. 122B:367-380(1999).
CC -1- FUNCTION: Initiates extracellular glutathione (GSH) breakdown, provides cells with a local cysteine supply and contributes to maintain intracellular GSH level. It is part of the cell

antioxidant defense mechanism. Catalyzes the transfer of the glutamyl moiety of glutathione to amino acids and dipeptide acceptors. Alternatively, glutathione can be hydrolyzed to give Cys-Gly and gamma glutamate.
CC -1- CATALYTIC ACTIVITY: (S-L-glutamyl)-peptide + an amino acid = peptide + 5-L-glutamyl-amino acid.
CC -1- PATHWAY: Plays a key role in the gamma-glutamyl cycle, a pathway for the synthesis and degradation of glutathione.
CC -1- SUBUNIT: Heterodimer composed of the light and heavy chains. The active site is located in the light chain (By similarity).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (By similarity).
CC -1- TISSUE SPECIFICITY: Highly expressed in kidney. Detected at lower levels in liver, lung, plexus chorioideus and brain capillary endothelial cells.
CC -1- SIMILARITY: Belongs to the gamma-glutamyltransferase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; Z46922; CAA87031.1; -
EMBL; X16533; CAA34536.1; -
PIR; S05532; S05532.
MEROPS; T03.006; -
InterPro; IPR000101; Peptidase_T3.
PIfam; PF01019; G_glu_transpept; 1.
PRINTS; PR01210; GGTTRANSPTASE.
TIGRFAMs; TIGR00065; G_glu_trans; 1.
PROSITE; PS00462; G_GLU_TRANSPEPTIDASE; 1.
Glutathione biosynthesis; Transferase; Acyltransferase; Signal-anchor; Transmembrane; Zymogen; Glycoprotein.
FT CHAIN 1 379 GAMMA-GLUTAMYLTRANSPEPTIDASE 1 HEAVY CHAIN.
FT CHAIN 380 568 GAMMA-GLUTAMYLTRANSPEPTIDASE 1 LIGHT CHAIN.
FT DOMAIN 1 4 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 5 26 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (PROBABLE).
FT DOMAIN 27 568 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 94 94 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 229 229 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 343 343 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 427 427 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 510 510 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 568 AA; 61315 MW; EB2D1896B1229487 CRC64;

Query Match 1.4%; Score 8; DB 1; Length 568;
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 348 RLAKTYET 355
DB 211 RLAKTYET 218

RESULT 19
PTI_ECOLI
ID_PTI_ECOLI STANDARD; PRT; 575 AA.
AC P08839;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Phosphoenolpyruvate-protein phosphotransferase (EC 2.7.3.9)
DE (Phosphotransferase system, enzyme 1).
GN PTSI OR B2416.
OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE=88314869; PubMed=2457575;
RX de Reuse H., Danchin A.;
RA "The ptsH, ptsI, and crr genes of the Escherichia coli
RT phosphoenolpyruvate-dependent phosphotransferase system: a complex
RT operon with several modes of transcription.";
RL J. Bacteriol. 170:3827-3837(1988).
RN [2]
RP SEQUENCE FROM N.A.
RP MEDLINE=88058992; PubMed=2960675;
RX Saffien D.W., Presper K.A., Doering T.L., Roseman S.;
RA "Sugar transport by the bacterial phosphotransferase system.
RT Molecular cloning and structural analysis of the Escherichia coli
RT ptsH, ptsI, and crr genes.";
RL J. Biol. Chem. 262:16241-16253(1987).
RN [3]
RP SEQUENCE FROM N.A.
RP STRAIN=K12 / MG1655;
RX MEDLINE=8257033; PubMed=9278503;
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glaesner J.D., Rode C.K., Mayhew G.P.,
RA Gregory J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RP STRAIN=K12;
RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Alba T., Baba T., Hayaashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitohashi N.,
RA Mirobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA Oshima T., Oyama S., Saito N., Sempel G., Satoh Y., Sivasubram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT -K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features.";
RL DNA Res. 4:191-113(1997).
RN [5]
RP SEQUENCE OF 1-56 FROM N.A.
RX MEDLINE=85286351; PubMed=2411636;
RA de Reuse H., Roy A., Danchin A.;
RT "Analysis of the ptsH-ptsI-crr region in Escherichia coli K-12:
RT nucleotide sequence of the ptsH gene.";
RL Gene 35:199-207(1985).
RN [6]
RP SEQUENCE OF 1-54 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=88257033; PubMed=3290198;
RA Byrne C.R., Monroe R.S., Ward K.A., Kredich N.M.;
RT "DNA sequences of the cysK regions of Salmonella typhimurium and
RT Escherichia coli and linkage of the cysK regions to ptsH.";
RL J. Bacteriol. 170:3150-3157(1988).
RN [7]
RP SEQUENCE OF 1-11.
RC STRAIN=K12 / BWG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RT in the genome of Escherichia coli K-12.";
RL Electrophoresis 18:1259-1313(1997).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 1-259.
RX MEDLINE=96434331; PubMed=8805571;
RA Liao D.-I., Silverton E., Seok Y.-J., Lee B.R., Peterkofsky A.,
RA Davies D.R.;
RT "The first step in sugar transport: crystal structure of the amino
RT terminal domain of enzyme I of the E. coli PEP: sugar

phosphotransferase system and a model of the phosphotransfer complex
with HPr.";
Structure 4:861-872(1996).
[9]
RP STRUCTURE BY NMR OF 1-259.
RP MEDLINE=97207064; PubMed=9054557;
RX Garrett D.S., Seok Y.-J., Liao D.-I., Peterkofsky A., Gronenborn A.M.,
RA Clore G.M.;
RT "Solution structure of the 30 kDa N-terminal domain of enzyme I of
RT the Escherichia coli phosphoenolpyruvate:sugar phosphotransferase
RT system by multidimensional NMR.";
RL Biochemistry 36:2517-2530(1997).
RN [10]
RP STRUCTURE BY NMR OF 1-259.
RP MEDLINE=98200485; PubMed=9541412;
RX Garrett D.S., Seok Y.-J., Peterkofsky A., Clore G.M., Gronenborn A.M.;
RA "Tautomeric state and pKa of the phosphorylated active site histidine
RT in the N-terminal domain of enzyme I of the Escherichia coli
RT phosphoenolpyruvate:sugar phosphotransferase system.";
RL Protein Sci. 7:789-793(1998).
RN [11]
RP STRUCTURE BY NMR OF 1-259.
RP MEDLINE=99140298; PubMed=10048929;
RX Garrett D.S., Seok Y.-J., Peterkofsky A., Gronenborn A.M., Clore G.M.;
RA "Solution structure of the 40,000 Mr phosphoryl transfer complex
RT between the N-terminal domain of enzyme I and HPr.";
RL Nat. Struct. Biol. 6:166-173(1999).
CC -1- FUNCTION: This is a component of the phosphoenolpyruvate-dependent
CC sugar phosphotransferase system (PTS), a major carbohydrate active
CC -transport system. Enzyme I transfers the phosphoryl group from
CC phosphoenolpyruvate (PEP) to the phosphoryl carrier protein (HPr).
CC Enzyme I is common to all PTS.
CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + protein L-histidine =
CC pyruvate + protein N(pi)-phospho-L-histidine.
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the PEP-utilizing enzyme family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; J02796; AAA24441.1; -
EMBL; M10425; AAA24439.1; -
EMBL; AE000329; AAC75469.1; -
EMBL; D90871; BAA15290.1; -
EMBL; M21994; BAA24385.1; -
EMBL; M21451; AAA23656.1; -
PIR; E29785; WQSCP1.
DR PDB; 1EZA; 07-JAN-98.
DR PDB; 2EZA; 20-AUG-97.
DR PDB; 1EZB; 07-JAN-98.
DR PDB; 2EZB; 20-AUG-97.
DR PDB; 1E2C; 07-JAN-98.
DR PDB; 2E2C; 20-AUG-97.
DR PDB; 1E2D; 07-JAN-98.
DR PDB; 3E2B; 29-DEC-99.
DR PDB; 3E2E; 16-DEC-98.
DR PDB; 12VM; 07-DEC-96.
DR PDB; 3EZA; 25-MAY-99.
DR SWISS-PROT; P08833; COLI.
DR EC02DBASE; B058.3; 6TH EDITION.
DR EcoGene; E010789; ptsI...
DR InterPro; IPR008731; PEP-utilisers_N.
DR InterPro; IPR008279; PEP mobile.
DR InterPro; IPR006318; PEP_P_trans.
DR InterPro; IPR000121; PEP-utilizers.
DR Pfam; PF05524; PEP-utilisers_N; 1.
DR Pfam; PF00391; PEP-utilizers; 1.

DR Pfam: PF02896; PEP-utilizers C: 1.
DR PRINTS: PR01736; PHPTNFRASE
DR ProDom: PD000940; PEP-utilizers; 1.
DR TIGRFAMs: TIGR04417; PIS_1fam; 1.
DR PROSITE: PS00370; PEP-ENZYMES_2; 1.
DR PROSITE: PS00742; PEP-ENZYMES_2; 1.
KW Phosphotransferase system; Transferase; Kinase; Sugar transport;
FT Phosphorylation; 3D-structure; Complete proteome.
FT ACT_SITE 189 189
FT MOD_RES 189 189
FT STRAND 8 8
FT STRAND 11 18
FT STRAND 30 30
FT STRAND 33 35
FT TURN 33 35
FT HELIX 36 63
FT TURN 64 65
FT TURN 66 80
FT HELIX 66 80
FT TURN 81 81
FT HELIX 83 94
FT TURN 95 96
FT STRAND 98 98
FT HELIX 100 115
FT TURN 116 117
FT HELIX 121 142
FT TURN 143 143
FT HELIX 149 151
FT STRAND 156 159
FT HELIX 165 168
FT TURN 169 170
FT TURN 173 175
FT STRAND 176 180
FT HELIX 189 197
FT TURN 198 198
FT STRAND 201 202
FT HELIX 208 210
FT TURN 211 211
FT TURN 214 215
FT STRAND 217 221
FT TURN 222 225
FT STRAND 226 229
FT HELIX 233 240
FT TURN 241 246
SQ SEQUENCE 575 AA; 63561 MW; 4278F0838855E950 CRC64;
Query Match 1.4%; Score 8; DB 1; Length 575;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 184 ELRDEGKA 191
Db 412 ELRDEGKA 419
RESULT 20
FT -SALTY
ID FT1 -SALTY STANDARD; PRT; 575 AA.
AC F12654;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Phosphoenolpyruvate-protein phosphotransferase (EC 2.7.3.9)
GN (Phosphotransferase system, enzyme I).
OS PSI OR STM2432 OR STY2668 OR T0425.
OS Salmonella typhimurium, and
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium;
RX MEDLINE=92011751; PubMed=1655788;
RA Licalisi C., Crocenzi T.S., Freire E., Roseman S.;
"Sugar transport by the bacterial phosphotransferase system. Structural and thermodynamic domains of enzyme I of Salmonella typhimurium"; J. Biol. Chem. 266:19519-19527(1991).
[2]
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P., Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D., Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E., Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M., Waterston K., Wilson R.K.; of Salmonella enterica serovar Typhimurium LT2."; Nature 413:852-856(2001).
[3]
RN [3]
RP SEQUENCE OF 1-299 FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2;
RX MEDLINE=86257033; PubMed=3290198;
RA Byrne C.R., Monroe R.S., Ward K.A., Kredich N.M.;
"DNA sequences of the cysK regions of Salmonella typhimurium and Escherichia coli and linkage of the cysK regions to ptsH."; J. Bacteriol. 170:3150-3157(1988).
[4]
RN [4]
RP SEQUENCE OF 1-9 FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2;
RX MEDLINE=89237892; PubMed=2497295;
RA Schnierow B.J., Yamada M., Saier M.H. Jr.;
"Partial nucleotide sequence of the pts operon in Salmonella typhimurium: comparative analyses in five bacterial genera."; Mol. Microbiol. 3:113-118(1989).
[5]
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J., Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M., Baker S., Basham D., Brooks K.M., Dowd L., White N., Farrar J., Cronin A., Davis P., Davies R.M., Hien T.T., Holroyd S., Jagels K., Krogsh A., Larsen T.S., Leather S., Moulse S., O'Gaora P., Parry C., Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.;
"Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18."; Nature 413:848-852(2001).
[6]
RN [6]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J., Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
"Comparative genomics of Salmonella enterica serovar Typhi strains Ty2 and CT18."; J. Bacteriol. 185:2330-2337(2003).
-!- FUNCTION: This is a component of the phosphoenolpyruvate-dependent sugar phosphotransferase system (PTS), a major carbohydrate active -transport system. Enzyme I transfers the phosphoryl group from phosphoenolpyruvate (PEP) to the phosphoryl carrier protein (HPr). Enzyme I is common to all PTS.
-!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + protein L-histidine = pyruvate + protein N(pi)-phospho-L-histidine.
-!- SUBUNIT: Homodimer.
-!- SUBCELLULAR LOCATION: Cytoplasmic.
-!- SIMILARITY: Belongs to the PEP-utilizing enzyme family.
-!- CAUTION: In strain CT18 it seems to be a pseudogene. It is interrupted by a frameshift in position 353. The sequence has been verified by the authors and is believed to be correct.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its

```

-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
EMBL; M38272; AAA73520.1; -.
PIR; A37970; FPGO.
HSP; P02768; 1E7B.
InterPro; IPR000264; Serum albumin.
Pfam; PF00273; transport Prot; 3.
PRINTS; PR00802; SERUMALBUMIN.
ProDom; PD002486; Serum albumin; 1.
SMART; SM00103; ALBUMIN; 3.
PROSITE; PS00212; ALBUMIN; 2.
Glycoprotein; Sulfation; Repeat; Metal-binding; Copper; Nickel;
KW      SIGNAL.
FT      1      18      BY SIMILARITY.
FT      CHAIN      19      609      ALPHA-FETOPROTEIN.
FT      DOMAIN      20      205      ALBUMIN 1.
FT      DOMAIN      212      397      ALBUMIN 2.
FT      DOMAIN      404      595      ALBUMIN 3.
FT      METAL      22      22      COPPER AND NICKEL (BY SIMILARITY).
FT      DISULFID      99      114      BY SIMILARITY.
FT      DISULFID      113      124      BY SIMILARITY.
FT      DISULFID      148      193      BY SIMILARITY.
FT      DISULFID      192      201      BY SIMILARITY.
FT      DISULFID      224      270      BY SIMILARITY.
FT      DISULFID      269      277      BY SIMILARITY.
FT      DISULFID      289      303      BY SIMILARITY.
FT      DISULFID      302      313      BY SIMILARITY.
FT      DISULFID      384      393      BY SIMILARITY.
FT      DISULFID      416      462      BY SIMILARITY.
FT      DISULFID      461      472      BY SIMILARITY.
FT      DISULFID      485      501      BY SIMILARITY.
FT      DISULFID      500      511      BY SIMILARITY.
FT      DISULFID      538      583      BY SIMILARITY.
FT      CARBOHYD      582      591      BY SIMILARITY.
FT      SEQUENCE      609 AA; 68697 MW; E8AE548377DB60EB CRC64;
                                         N-LINKED (GLCNAC...) (POTENTIAL).
Query Match      1.4%; Score 8; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 10;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      140 YEIARRHP 147
Db      164 YEIARRHP 171
          |||||
          |||||

RESULT 22
PETA_HUMAN
ID      PETA_HUMAN      STANDARD;      PRT;      609 AA.
DT      21-JUL-1986 (Rel. 01, Created)
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Alpha-fetoprotein precursor (Alpha-fetoglobulin) (Alpha-1-
DE      fetoprotein).
DE      APP.
GN      Homo sapiens (Human).
OS      Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Eukaryota; Metazoa; Chordata;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=83273664; PubMed=6192439;
RA      Morinaga T., Sakai M., Wegmann T.G., Tamaoki T.;
RT      "Primary structures of human alpha-fetoprotein and its mRNA.";
RL      Proc. Natl. Acad. Sci. U.S.A. 80:4604-4608(1983).

```

[2] SEQUENCE FROM N.A.
RP MEDLINE=87185438; PubMed=2436661;
RX Gibbs P.E.M., Zielinski R., Boyd C., Dugaiczky A.;
RA "Structure, polymorphism, and novel repeated DNA elements revealed by
RT a complete sequence of the human alpha-fetoprotein gene.";
RL Biochemistry 26:1332-1343(1987).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abranson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 1-28 FROM N.A.
RX MEDLINE=93278385; PubMed=7684942;
RA McVey J.H., Michaelides K., Hansen L.P., Ferguson-Smith M.,
RA Tilghman S., Krumlauf R., Tuddenham E.G.D.;
RT "A G--A substitution in an HNF I binding site in the human alpha-
RT fetoprotein gene is associated with hereditary persistence of alpha-
RT fetoprotein (HPAFP).";
RL Hum. Mol. Genet. 2:379-379(1993).
RN [5]
RP SEQUENCE OF 429-556 FROM N.A.
RX MEDLINE=83158778; PubMed=6187626;
RA Beattie W.G., Dugaiczky A.;
RT "Structure and evolution of human alpha-fetoprotein deduced from
RT partial sequence of cloned cDNA.";
RL Gene 20:415-422(1982).
RN [6]
RP PARTIAL SEQUENCE OF 19-609.
RX MEDLINE=91242409; PubMed=1709810;
RA Pucci P., Siciliano R., Malorni A., Marino G., Tecce M.F.,
RA Ceccarini C., Terrana B.;
RT "Human alpha-fetoprotein primary structure: a mass spectrometric
RT study.";
RL Biochemistry 30:5061-5066(1991).
RN [7]
RP PRELIMINARY SEQUENCE OF 19-35.
RX MEDLINE=7742506; PubMed=70228;
RA Yachnin S., Hsu R., Heinrichson R.L., Miller J.B.;
RT "Studies on human alpha-fetoprotein. Isolation and characterization
RT of monomeric and polymeric forms and amino-terminal sequence
RT analysis.";
RL Biochim. Biophys. Acta 493:418-428(1977).
RN [8]
RP PRELIMINARY SEQUENCE OF 19-38.
RX MEDLINE=78001760; PubMed=711198;
RA Aoyagi Y., Ikenaka T., Ichida F.;
RT "Comparative chemical structures of human alpha-fetoproteins from
RT fetal serum and from ascites fluid of a patient with hepatoma.";
RL Cancer Res. 37:3663-3667(1977).
RN [9]
RP PRELIMINARY SEQUENCE OF 19-39.
RX MEDLINE=75018719; PubMed=4138095;
RA Ruoslahti E., Pihko H., Vaheri A., Seppala M., Virolainen M.,
RA Kontinen A.;
RT "Alpha fetoprotein: structure and expression in man and inbred mouse
RT strains under normal conditions and liver injury.";
RL Johns Hopkins Med. J. Suppl. 3:249-255(1974).
RN [10]
RP GENE STRUCTURE
RX MEDLINE=85182629; PubMed=2580930;
RA Sakai M., Morinaga T., Urano Y., Watanabe K., Wegmann T.G.,
RA Tamaoki T.;
RT "The human alpha-fetoprotein gene. Sequence organization and the 5'
RT flanking region.";
RL J. Biol. Chem. 260:5055-5060(1985).
RN [11]
RP METAL-BINDING.
RX MEDLINE=79001617; PubMed=80265;
RA Aoyagi Y., Ikenaka T., Ichida F.;
RT "Copper(II)-binding ability of human alpha-fetoprotein.";
RL Cancer Res. 38:3483-3486(1978).
RN [12]
RP BILIRUBIN-BINDING.
RX MEDLINE=80001710; PubMed=89900;
RA Aoyagi Y., Ikenaka T., Ichida F.;
RT "Alpha-Fetoprotein as a carrier protein in plasma and its bilirubin-
RT binding ability.";
RL Cancer Res. 39:3571-3574(1979).
RN [13]
RP SULFATION
RX MEDLINE=86042625; PubMed=2414772;
RA Liu M.C., Yu S., Sy J., Redman C.M., Lipmann F.;
RT "Tyrosine sulfation of proteins from the human hepatoma cell line
RT HepG2.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:7160-7164(1985).
RN [14]
RP FUNCTION: Binds copper, nickel, and fatty acids as well as, and
CC bilirubin less well than, serum albumin. Only a small percentage
CC (less than 2%) of the human AFP shows estrogen-binding properties.
CC -! SUBUNIT: Dimeric and trimeric forms have been found in addition
CC to the monomeric form.
CC -! SUBCELLULAR LOCATION: Secreted.
CC -! TISSUE SPECIFICITY: Plasma. Synthesized by the fetal liver and
CC yolk sac.
CC -! DEVELOPMENTAL STAGE: Occurs in the plasma of fetuses more than 4
CC weeks old, reaches the highest levels during the 12th-16th week of
CC gestation, and drops to trace amounts after birth. The serum level
CC in adults is usually less than 40 ng/ml. AFP occurs also at high
CC levels in the plasma and ascitic fluid of adults with hepatoma.
CC -! PTM: Independent studies suggest heterogeneity of the amino-
CC terminal sequence of the mature protein and of the cleavage site
CC of the signal sequence.
CC -! PTM: Sulfated.
CC -! SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC -! SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M10949; AAA51674.1; -
CC EMBL; M10950; AAA51675.1; -
CC EMBL; V01514; CAA24758.1; -
CC EMBL; M16110; AAB58754.1; -
CC EMBL; BC027861; AAB27861.1; -
CC EMBL; Z19532; CAA79592.1; -
CC PIR; A26624; FPHU.
CC HSSP; P02768; 1E7B.
CC GlycoSuiteDB; P02771; -
CC Sienna-2DPAGE; P02771; -
CC Genew; HGNC:317; AFP.
CC MIM; 104150; -


```
DR InterPro: IPR000264; Serum albumin.
DR Pfam: PF00273; transport_prot; 3.
DR PRINTS: PR00802; SERUMALBUMIN.
DR ProDom: PD002486; Serum_albumin; 1.
DR SMART: SM00103; ALBUMIN; 3.
DR PROSITE: PS00212; ALBUMIN; 2.
KW Glycoprotein; Sulfation; Repeat; Metal-binding; Copper; Nickel;
KW Signal; Polymorphism.
FT SIGNAL 1 18
FT CHAIN 19 609 ALPHA-FETOPROTEIN.
FT DOMAIN 20 205 ALBUMIN 1.
FT DOMAIN 212 397 ALBUMIN 2.
FT DOMAIN 404 595 ALBUMIN 3.
FT METAL 22 22 COPPER AND NICKEL.
FT DISULFID 99 114
FT DISULFID 113 124
FT DISULFID 148 193
FT DISULFID 192 201
FT DISULFID 224 270
FT DISULFID 269 277
FT DISULFID 289 303
FT DISULFID 302 313
FT DISULFID 384 393
FT DISULFID 416 462
FT DISULFID 461 472
FT DISULFID 485 501
FT DISULFID 500 511
FT DISULFID 538 583
FT DISULFID 582 591
FT CARBOHYD 251 251
FT VARIANT 570 570
FT SEQUENCE 609 AA; 68677 MW; 4D4E45820E1C2D4F CRC64;
SQ
Query Match 1.4%; Score 8; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 140 YEIARRHP 147
Db 164 YEIARRHP 171
|||||
|164 YEIARRHP 171|

RESULT 23
ID FETA_PANTR STANDARD; PRT; 609 AA.
AC Q28789;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Alpha-fetoprotein precursor (Alpha-fetoglobulin) (Alpha-1-
DE fetoprotein).
GN AFP.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96032345; PubMed=7557431;
RA Nishio H., Gibbs P.B., Minghetti P.P., Zielinski R., Dugaiczak A.;
RT "The chimpanzee alpha-fetoprotein-encoding gene shows structural
RT similarity to that of gorilla but distinct differences from that of
RT human.";
RL Gene 162:213-220(1995).
CC -!- FUNCTION: Binds copper, nickel, and fatty acids as well as, and
CC bilirubin less well than, serum albumin.
CC -!- SUBUNIT: Dimeric and trimeric forms have been found in addition
CC to the monomeric form (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma. Synthesized by the fetal liver and
CC yolk sac.
```

RA Caseady A. J., Salkild C.K., Baverstock P., Wallace J.C.;
 RL Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-28 FROM N.A.
 RX MEDLINE=83161037; PubMed=6187737;
 RA Hache R.J.G., Wiskocil R., Vasa M., Roy R.N., Lau P.C.K., Deesley R.G.;
 RL "the 5' noncoding and flanking regions of the avian very low density
 RT apolipoprotein II and serum albumin genes. Homologies with the egg
 RL white protein genes";
 RL J. Biol. Chem. 258:4556-4564 (1983).
 RN [3]
 RP SEQUENCE OF 19-30.
 RX MEDLINE=78019943; PubMed=911327;
 RA Rosen A.M., Geller D.M.;
 RL "Chicken microsomal albumin: amino terminal sequence of chicken
 RT proalbumin";
 RL Biochem. Biophys. Res. Commun. 78:1060-1066 (1977).
 RN [4]
 RP ALLERGENIC PROPERTIES.
 RX MEDLINE=21381307; PubMed=11488669;
 RA Quirce S., Maranon F., Umpierrez A., de las Heras M.,
 RA Fernandez-Caldas E., Sastre J.;
 RL "Chicken serum albumin (Gal d 5*) is a partially heat-labile inhalant
 RT and food allergen implicated in the bird-egg syndrome";
 RL Allergy 56:754-762 (2001).
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 CC hormones, bilirubin and drugs. Its main function is the regulation
 CC of the colloidal osmotic pressure of blood.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- ALLERGEN: Causes an allergic reaction in human. Binds IgE.
 CC Partially heat-labile allergen that may cause both respiratory and
 CC food-allergy symptoms in patients with the bird-egg syndrome.
 CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
 CC -!- SIMILARITY: Contains 3 albumin domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X60688; CAA43098.1; -;
 DR EMBL; V00381; CAA23680.1; -;
 DR PIR; S15571; ABCHS.
 DR HSSP; P02768; 1E7B.
 DR InterPro; IPR000264; Serum albumin.
 DR Pfam; PF00273; transport prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 3.
 DR Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
 KW SIGNAL 1 18
 FT PROPEP 19 23
 FT CHAIN 24 615 SERUM ALBUMIN.
 FT DOMAIN 24 209 ALBUMIN 1.
 FT DOMAIN 216 401 ALBUMIN 2.
 FT DOMAIN 408 599 ALBUMIN 3.
 FT METAL 30 30 COPPER (BY SIMILARITY).
 FT METAL 80 89 BY SIMILARITY.
 FT DISULFID 102 118 BY SIMILARITY.
 FT DISULFID 117 128 BY SIMILARITY.
 FT DISULFID 152 197 BY SIMILARITY.
 FT DISULFID 196 205 BY SIMILARITY.
 FT DISULFID 228 274 BY SIMILARITY.
 FT DISULFID 273 281 BY SIMILARITY.
 FT DISULFID 293 307 BY SIMILARITY.
 FT DISULFID 306 317 BY SIMILARITY.
 FT DISULFID 344 389 BY SIMILARITY.

FT DISULFID 388 397 BY SIMILARITY.
 FT DISULFID 420 476 BY SIMILARITY.
 FT DISULFID 465 476 BY SIMILARITY.
 FT DISULFID 489 505 BY SIMILARITY.
 FT DISULFID 515 515 BY SIMILARITY.
 FT DISULFID 542 587 BY SIMILARITY.
 FT DISULFID 586 595 BY SIMILARITY.
 FT CARBOHYD 500 500 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 24 24 F -> M (IN REF. 3).
 SQ SEQUENCE 615 AA; 59918 MW; E59E4B8C8C066C6 CRC64;
 Query Match 1.4%; Score 8; DB 1; Length 615;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 96 PERNECFL 103
 DB 123 PERNECFL 130
 RESULT 25
 EXSB BORBU STANDARD; PRT; 1169 AA.
 AC O51578;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Exodeoxyribonuclease V beta chain (EC 3.1.11.5).
 GN RECB OR B06033.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A. / B31;
 RC STRAIN=ATCC 35210 / B31;
 RX MEDLINE=98065943; PubMed=9403685;
 RA Fraser C.M., Carjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
 RA Petherston J., Tomb J.-F., Fleischmann R.D., Richardson D.,
 RA Peterson B., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
 RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
 RA Utterback T., Matthey L., McDonald L., Artiach P., Bowman C.,
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
 RA Smith H.O., Venter J.C.;
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
 RT burgdorferi";
 RL Nature 390:580-586 (1997).
 CC -!- FUNCTION: REQUIRED FOR EFFICIENT DNA REPAIR; IT CATALYZES THE
 CC UNWINDING OF DOUBLE-STRANDED DNA AND THE CLEAVAGE OF SINGLE-
 CC STRANDED DNA AND IT STIMULATES LOCAL GENETIC RECOMINATION.
 CC ALL OF THESE ACTIVITIES REQUIRE CONCOMITANT HYDROLYSIS OF ATP
 CC (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage (in the presence of
 CC ATP) in either 5' to 3' or 3' to 5'-direction to yield 5'-
 CC phosphooligonucleotides.
 CC -!- SUBUNIT: Consist of three subunits; recB, recC and recD (By
 CC similarity).
 CC -!- SIMILARITY: Belongs to the helicase family. UvrD subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE001164; AAC66981.1; -;
 DR PIR; H70178; H70178.
 DR HSSP; P56255; 1PJR.
 DR TIGR; B50633; -;
 DR InterPro; IPR0004586; RecB.
 DR InterPro; IPR000212; UvrD-helicase.

DR Pfam: PF00580; UvrD-helicase; 1.
 DR TIGRFAMs; TIGR00609; recB; 1.
 KW Hydrolyase; Nuclease; Exonuclease; Endonuclease; Helicase; ATP-binding;
 KW DNA repair; Complete proteome.
 FT BIND 18 25 ATP (POTENTIAL).
 SQ SEQUENCE 1169 AA; 137828 MW; B61D63C1C959B91F CRC64;

Query Match
 Best Local Similarity 1.4%; Score 8; DB 1; Length 1169;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 351 KTYETTL 358
 |||||
 DB 684 KTYETTL 691

RESULT 26
 UL11_HSV6U STANDARD; PRT; 77 AA.
 AC P24448;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein U71.
 GN U71 OR 17R.
 OS Human herpesvirus (type 6 / strain Uganda-1102) (HHV6).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Roseolovirus.
 OX NCBI_TaxID=10370;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90080132; PubMed=2152817;
 RA Lawrence G.L.; Chee M.; Craxton M.A.; Gompels U.A.; Honess R.W.;
 RA Barrell B.G.;
 RT "Human herpesvirus 6 is closely related to human cytomegalovirus";
 RL J. Virol. 64:287-299 (1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9526321; PubMed=7747482;
 RA Gompels U.A.; Nicholas J.; Lawrence G.; Jones M.; Thomson B.J.;
 RA Martin M.E.; Efstathiou S.; Craxton M.; Macaulay H.A.;
 RT "The DNA sequence of human herpesvirus-6: structure, coding content,
 RT and genome evolution";
 RL Virology 209:29-51 (1995).
 CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL11,
 CC HSV-1 51, HSV-6 ORF17R, EBV BFLP1, AND VZV 49.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

EMBL; X83413; CAA58363.1; -
 DR EMBL; M68963; AAA65579.1; -
 DR PIR; G36769; G36769.
 KW Hypothetical protein.
 SQ SEQUENCE 77 AA; 8469 MW; B58CF2D7637CDB02 CRC64;

Query Match
 Best Local Similarity 1.2%; Score 7; DB 1; Length 77;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 EKPLEK 286
 |||||
 DB 52 EKPLEK 58

RESULT 27
 ATPE_GALSU STANDARD; PRT; 138 AA.
 ID _ATPE_GALSU

AC Q08908;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE ATP synthase epsilon chain (EC 3.6.3.14) (ATP synthase F1 sector
 DE epsilon subunit).
 GN ATP8.
 OS Gaillardia sulphuraria (Red alga).
 OC Chloroplast.
 OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;
 OC Gaillardia.
 OX NCBI_TaxID=130081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=14-1-1 / Isolate 107.79/Goettingen;
 RX MEDLINE=94033298; PubMed=8219057;
 RA Kozrzewa M.; Zetsche K.;
 RT "Organization of plastid-encoded ATPase genes and flanking regions
 RT including homologues of infB and tsf in the thermophilic red alga
 RT Gaillardia sulphuraria";
 RL Plant Mol. Biol. 23:67-76 (1993).
 CC -!- FUNCTION: Produces ATP from ADP in the presence of a proton
 CC gradient across the membrane.
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
 CC H(+) (Out).
 CC -!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic
 CC core - and CF(0) - the membrane proton channel. CF(1) has five
 CC subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0)
 CC has three main subunits: a, b and c.
 CC -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
 CC -!- SIMILARITY: Belongs to the ATPase epsilon chain family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

EMBL; X6698; CAA47242.1; -
 DR HAMAP; MF_00530; -; 1.
 DR InterPro; IPR001469; ATPsynt_DE.
 DR Pfam; PF00401; ATP-synt DE; 1.
 DR Pfam; PF02823; ATP-synt DE N; 1.
 DR ProDom; PD000944; ATPsynt_DE; 1.
 DR TIGRFAMs; TIGR01216; ATP_synt_epsi; 1.
 KW Hydrolyase; ATP synthetase; CF(1); Hydrogen ion transport; Thylakoid;
 KW Membrane; Chloroplast.
 SQ SEQUENCE 138 AA; 15057 MW; CB54B86BDBCBA62A CRC64;

Query Match
 Best Local Similarity 1.2%; Score 7; DB 1; Length 138;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 291 AVEVDE 297
 |||||
 DB 67 AVEVDE 73

RESULT 28
 CX44_FIG STANDARD; PRT; 138 AA.
 ID _CX44_FIG
 AC Q29559;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Gap junction alpha-4 protein (Connexin 37) (Cx37) (fragment).
 GN GJA4.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;

RN SEQUENCE FROM N.A.
 RP STRAIN=Large white X Duroc; TISSUE=Aortic endothelium;
 RX MEDLINE=96429298; PubMed=8832399;
 RA Carter T.D., Cen X.Y., Carille G., Kalapothakis E., Ogden D.,
 RA Evans W.H.;
 RA "Porcine aortic endothelial gap junctions: identification and
 RT permeation by caged InaP3";
 RL J. Cell Sci. 109:1765-1773(1996).
 CC -!- FUNCTION: One gap junction consists of a cluster of closely packed
 CC pairs of transmembrane channels, the connexons, through which
 CC materials of low MW diffuse from one cell to a neighboring cell.
 CC -!- SUBUNIT: A connexon is composed of a hexamer of connexins.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to the connexin family. Alpha-type (group II)
 CC subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC
 CC EMBL; X86024; CAA60019.1; -;
 DR InterPro; IPR000500; Connexin.
 DR Pfam; PF00029; connexin; 1.
 DR PRINTS; PR00026; CONNEXIN.
 DR SMART; SM00037; CNX; 1.
 DR PROSITE; PS00407; CONNEXINS 1; 1.
 DR PROSITE; PS00408; CONNEXINS 2; PARTIAL.
 KW Gap junction; Transmembrane.
 FT NON_TER 1 1
 FT DOMAIN <1 16
 FT TRANSMEM 17 39
 FT DOMAIN 40 74
 FT TRANSMEM 75 97
 FT DOMAIN 98 >138
 FT NON_TER 138 138
 FT SEQUENCE 138 AA; 15836 MW; 929886C6784C222 CRC64;
 SQ
 Query Match 1.2%; Score 7; DB 1; Length 138;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 418 VSTPTLV 424
 DB 83 VSTPTLV 89
 RESULT 29
 NPC2 MOUSE
 ID NPC2 MOUSE STANDARD; PRT; 149 AA.
 AC Q9Z0J0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Epidermal secretory protein El precursor.
 GN NPC2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Matsumoto K.;
 RA "Mus musculus mRNA for epidermal secretory protein.";
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Small intestine, and Tongue;

RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakawa I.,
 RA Saito T., Okazaki Y., Gojohri T., Bono H., Kasukawa T., Saito R.,
 RA Kado K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Giess C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
 RA Schraml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Wittaker C., Wilming L.,
 RA Wyshak-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kotsuki S.,
 RA Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection.";
 RT Nature 409:685-690(2001).
 RL
 RN
 RP
 RQ
 SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary gland;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Faney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RL
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the NPC2 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC
 CC EMBL; AB021289; BAA35183.1; -;
 DR EMBL; AK009127; BAB26090.1; -;
 DR EMBL; AK008603; BAB25771.1; -;
 DR EMBL; BC003471; AAB03471.1; -;
 DR EMBL; BC007190; AAB07190.1; -;
 DR MGD; MGI:1915213; Npc2.
 DR InterPro; IPR003172; El_DerP2_DerF2.
 DR InterPro; IPR007110; IG-like.
 DR Pfam; PF02221; El_DerP2_DerF2; 1.
 DR SMART; SM00737; NF; 1.
 DR Glycoprotein; Signal;
 KW SIGNAL 1 19
 FT CHAIN 20 149
 FT DISULFID 27 140
 FT DISULFID 42 47
 FT DISULFID 93 99
 FT CARBOHYD 58 58
 FT POTENTIAL.
 FT EPIDIDYMAL SECRETORY PROTEIN El.
 FT BY SIMILARITY.
 FT BY SIMILARITY.
 FT BY SIMILARITY.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 69 69 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 149 AA; 16442 MW; 68DE36CF69791805 CRC64;

Query Match 1.2%; Score 7; DB 1; Length 149;
Best Local Similarity 100.0%; Pred.No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 575 LVAASQA 581
DB 13 LVAASQA 19

RESULT 30
YF56 MYCTU STANDARD; PRT; 202 AA.
AC Q10774;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical transcriptional regulator Rv1556/MT1607/MD1581.
GN Rv1556 OR MT1607 OR MTCV48.09C OR MB1581.
OS Mycobacterium tuberculosis, and
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773, 1765;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whithead S., Barrall B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL J. Bacteriol. 184:5479-5490(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=M.bovis; STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12789972;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthey S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; Z74020; CAA98307.1; -.
DR EMBL; AE007026; AAK45874.1; -.
DR EMBL; BX248339; CAD96248.1; -.
DR PIR; A70763; A70763.
DR TIGR; MT1607; -.
DR TubercuList; Rv1556; -.
DR InterPro; IPR001647; HTH_Tetr.
DR Pfam; PF00440; tetr; 1.
DR PRINTS; PR00455; HTHTETR.
DR PROSITE; PS01081; HTH_TETR_FAMILY; FALSE_NEG.
KW Hypothetical protein; Transcription regulation; DNA-binding;
KW Complete proteome.
SQ SEQUENCE 202 AA; 22329 MW; 6754AD8609195C42 CRC64;

Query Match 1.2%; Score 7; DB 1; Length 202;
Best Local Similarity 100.0%; Pred.No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 423 LVEVSRN 429
DB 114 LVEVSRN 120

RESULT 31
RS4_RICCN STANDARD; PRT; 205 AA.
ID RS4_RICCN
AC Q92IF2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S4.
GN RPSD OR RC0468.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Malish 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samson D., Roux S., Cossart P., Weissenbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii."
RL Science 293:2093-2098(2001).
CC -!- FUNCTION: One of the primary rRNA binding proteins, it binds
CC directly to 16S rRNA where it nucleates assembly of the body of
CC the 30S subunit (By similarity).
CC -!- FUNCTION: With S5 and S12 plays an important role in translational
CC accuracy (By similarity).
CC -!- SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.
CC The interaction surface between S4 and S5 is involved in control
CC of translational fidelity (By similarity).
CC -!- SIMILARITY: Contains 1 S4 RNA-binding domain.
CC -!- SIMILARITY: Belongs to the S4P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AE008610; AAL03006.1; -.
DR PIR; D97758; D97758.
DR HAMAP; MF_01306; -.
DR InterPro; IPR001912; Ribosomal_S4.
DR InterPro; IPR005709; Ribosomal_S4_b/o.
DR InterPro; IPR002942; S4.
DR Pfam; PF00163; Ribosomal_S4; 1.
DR Pfam; PF01479; S4; 1.

DR SMART; SM00363; S4; 1.
 DR TIGRFAMS; TIGR01017; rpsd bact; 1.
 DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
 DR PROSITE; PS00889; S4; 1.
 KW Ribosomal protein; rRNA-binding; Complete proteome.
 FT DOMAIN 94 AA; 23136 MW; S4 RNA-BINDING.
 SQ SEQUENCE 205 AA; 23136 MW; CS0B753BA3EEDOC CRC64;

Query Match 1.2%; Score 7; DB 1; Length 205;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 AKORLKC 200
 DB 54 AKORLKC 60

RESULT 32
 RS4_RICPR STANDARD; PRT; 205 AA.
 AC Q9ZDI3,
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 30S ribosomal protein S4.
 GN RPSD OR RP345.
 OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsia; Rickettsia.
 OX NCBI_TaxID=782;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Madrid E;
 RX MEDLINE=99039499; PubMed=9823893;
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
 RA Sichteritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of Rickettsia prowazekii and the origin of
 RT mitochondria."
 RL Nature 396:133-140 (1998).
 CC -!- FUNCTION: One of the primary rRNA binding proteins, it binds
 CC directly to 16S rRNA where it nucleates assembly of the body of
 CC the 30S subunit (By similarity).
 CC -!- FUNCTION: With S5 and S12 plays an important role in translational
 CC accuracy (By similarity).
 CC -!- SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.
 CC of translational fidelity (By similarity).
 CC The interaction surface between S4 and S5 is involved in control
 CC of translational fidelity (By similarity).
 CC -!- SIMILARITY: Contains 1 S4 RNA-binding domain.
 CC -!- SIMILARITY: Belongs to the S4P family of ribosomal proteins.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AJ235271; CAAL4805.1; -.
 CC F01; C71691; C71691.
 CC HSP; P81288; 1C05.
 CC HAMAP; MF_01306; -; 1.
 CC InterPro; IPR001912; Ribosomal_S4.
 CC InterPro; IPR005709; Ribosomal_S4_b/o.
 CC InterPro; IPR002942; S4.
 CC Pfam; PF00163; Ribosomal_S4; 1.
 CC Pfam; PF01479; S4; 1.
 CC SMART; SM00363; S4; 1.
 CC TIGRFAMS; TIGR01017; rpsd bact; 1.
 CC PROSITE; PS00632; RIBOSOMAL_S4; 1.
 CC PROSITE; PS00889; S4; 1.
 KW Ribosomal protein; RNA-binding; Complete proteome.

FT DOMAIN 94 157 S4 RNA-BINDING.
 SQ SEQUENCE 205 AA; 23318 MW; 3750C8D28A9FAF49 CRC64;

Query Match 1.2%; Score 7; DB 1; Length 205;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 AKORLKC 200
 DB 54 AKORLKC 60

RESULT 33
 XKDN_BACSU STANDARD; PRT; 208 AA.
 ID XKDN_BACSU
 AC P54333;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Phage-like element PBX protein xkdn.
 GN XKDN OR ESU12670.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Boursier R., Boursier L., Brans A., Braun M., Brignall S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Giuseppe G., Guy B.J., Hega K., Haech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan B., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivoita C., Rocha E., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tanakashi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis".
 RL Nature 390:249-256 (1997).
 CC -!- SIMILARITY: STRONG, TO B.SUBTILIS YQBN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

```

CC
EMBL; Z70177; CAA94036.1; -
DR
EMBL; Z99110; CAB13124.1; -
DR
PIR; E69732; E69732.
DR
Subtilist; BG11548; xkdn.
KW
Complete proteome.
SQ
SEQUENCE 208 AA; 24081 MW; 8B461779CAAD4C6A CRC64;

Query Match 1.2%; Score 7; DB 1; Length 208;
Best Local Similarity 100.0%; Pred.No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 570 EBGKLV 576
|||||
DB 144 EBGKLV 150

RESULT 34
RRNH2 SYNFX
AC O7U4C6; STANDARD; PRT; 224 AA.
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ribonuclease HII (EC 3.1.26.4) (RNase HII).
RN RNBH OR SYNW2144.
OS Synecococcus sp. (strain WH8102).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=84598;
[1]
SEQUENCE FROM N.A.
RP MEDLINE=22825697; PubMed=12917641;
RX Palenik B., Brahmasha B., Larimer F.W., Land M., Hauser L., Chain P.,
RA Lamerdin J., Regala W., Allen E.E., McCarren J., Paulsen I.,
RA Dufresne A., Partensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synecococcus.";
RL Nature 424:1037-1042(2003).
CC -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of
CC RNA-DNA hybrids specifically (By similarity).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- COFACTOR: Manganese (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: Belongs to the RNase HII family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
EMBL; BX569694; CAE08659.1; ALT_INIT.
DR
HAMAP; MF_00052; -; 1
DR InterPro; IPR001352; RNase HII/HII.
DR Pfam; PF01351; RNase HII; 1.
DR Kwd Hydrolase; Nuclease; Endonuclease; Complete proteome.
FW ACT_SITE 42 42 BY SIMILARITY.
FT ACT_SITE 138 138 BY SIMILARITY.
FT ACT_SITE 157 157 BY SIMILARITY.
SQ SEQUENCE 224 AA; 24310 MW; 7F21360ABC4C54AB CRC64;

Query Match 1.2%; Score 7; DB 1; Length 224;
Best Local Similarity 100.0%; Pred.No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 218 RLSQFF 224
|||||
DB 180 RLSQFF 186

RESULT 35

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AF003139; AAB54165.1; -;
DR PIR; T29034; T29034.
DR WormPep; F53G12.10; CE11024.
DR InterPro; IPR000517; Ribosomal_L30.
DR InterPro; IPR005998; Ribosomal_L7_euk.
DR Pfam; PF00327; Ribosomal_L30; 1.
DR TIGRFAMS; TIGR01310; L7; 1.
DR PROSITE; PS00634; RIBOSOMAL_L30; FALSE_NEG.
KW Ribosomal protein; RNA-binding.
SQ SEQUENCE 244 AA; 28132 MW; 5073D697E90A8257 CRC64;

Query Match 1.2%; Score 7; DB 1; Length 244;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 412 TKKVPQV 418
| | | | |
DB 4 TKKVPQV 10

RESULT 37
YL22_YEAST STANDARD; PRT; 250 AA.
ID YL22_YEAST
AC Q07953;
DT 01-NOV-1997 (Rel. 35; Created)
DT 01-NOV-1997 (Rel. 35; Last sequence update)
DT 10-OCT-2003 (Rel. 42; Last annotation update)
DE Hypothetical UPF0023 protein YLR022C.
GN YLR022C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313267; PubMed=9169871;
RA Johnston M., Hillier L., Riles L., Albertmann K., Andre B., Ansorge W.,
RA Benes V., Brueckner M., Dubois E., Duesterhoeft A.,
RA Entian K.-D., Floeth M., Goffeau A., Heblung U., Heumann K.,
RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
RA Louis E.J., Messenguy P., Mewes H.-W., Miosga T., Moestl D.,
RA Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA Portetle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA Scharte M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Utrretarazu L.A., Vandenbol M., Verhasselt P.,
RA Vierdeels F., Voet M., Volckaert G., Voss H., Wambutt R., Weidner E.,
RA Wedler H., Zimmermann F.K., Zollner A., Hani J., Hohseisel J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL Nature 387:97-90 (1997).
CC -!- SIMILARITY: Belongs to the UPF0023 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; Z73194; CAA97545.1; -;
DR PIR; S64849; S64849.
DR Geronline; 142084; -;
DR SGD; S0004012; YLR022C.
DR InterPro; IPR002140; UPF0023.

DR Pfam; PF01172; UPF0023; 1.
DR ProDom; PD009796; UPF0023; 1.
DR TIGRFAMS; TIGR00291; TIGR00291; 1.
DR PROSITE; PS01267; UPF0023; 1.
KW Hypothetical protein.
SQ SEQUENCE 250 AA; 28283 MW; EFC799ADEFF73E0E CRC64;

Query Match 1.2%; Score 7; DB 1; Length 250;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 516 LSEKQ 522
| | | | |
DB 95 LSEKQ 101

RESULT 38
FR12_TOBAC STANDARD; PRT; 259 AA.
ID FR12_TOBAC
AC QH1T3;
DT 10-OCT-2003 (Rel. 42; Created)
DT 10-OCT-2003 (Rel. 42; Last sequence update)
DT 10-OCT-2003 (Rel. 42; Last annotation update)
DE Ferritin 2, chloroplast precursor (NtPer2).
GN FER2.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RA Jiang T., Yoshinara T., Masuda T., Goto F.;
RT "Occurrence and expression analysis of two types of the tobacco
RT ferritin genes.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Ferritin is an intracellular molecule that stores iron
CC in a soluble, nontoxic, readily available form. The functional
CC molecule, which is composed of 24 chains, is roughly spherical and
CC contains a central cavity into which the polymeric ferric iron
CC core is deposited (By similarity). (By similarity).
CC -!- SUBCELLULAR LOCATION: Chloroplast. (By similarity).
CC -!- SIMILARITY: Belongs to the ferritin family.
CC -!- SIMILARITY: Contains 1 ferritin-like diiron domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

EMBL; AY141105; AAN06322.1; -;
DR InterPro; IPR001519; Ferritin.
DR InterPro; IPR008331; Ferritin_Dps.
DR InterPro; IPR009040; Ferritin_Like.
DR Pfam; PF00210; ferritin; 1.
DR ProDom; PD000971; Ferritin; 1.
DR PROSITE; PS00540; FERRITIN_1; FALSE_NEG.
DR PROSITE; PS00204; FERRITIN_2; FALSE_NEG.
DR PROSITE; PS00905; FERRITIN_Like; 1.
KW Iron storage; Iron; Metal-Binding; Chloroplast; Transit peptide.
FT TRANSIT 1 52 CHLOROPLAST (POTENTIAL).
FT CHAIN 53 259 FERRITIN 2.
FT DOMAIN 53 85 EXTENSION PEPTIDE (EP).
FT DOMAIN 86 239 FERRITIN-LIKE DIIRON.
FT METAL 103 107 IRON (BY SIMILARITY).
FT METAL 137 137 IRON (BY SIMILARITY).
FT METAL 138 138 IRON (BY SIMILARITY).
FT METAL 140 140 IRON (BY SIMILARITY).
FT METAL 141 141 IRON (BY SIMILARITY).
SQ SEQUENCE 259 AA; 29220 MW; 59997A45142375C3 CRC64;


```
Query Match 1.2%; Score 7; DB 1; Length 259;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 307 ADFVESK 313
DB 210 ADFVESK 216

RESULT 39
GTH2 WHEAT
ID GTH2 WHEAT STANDARD; PRT; 291 AA.
AC P30111;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Glutathione S-transferase 2 (EC 2.5.1.18) (GST class-phi).
GN GSTA2.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Cheyenne;
RX MEDLINE=91322503; PubMed=1650615;
RA Mauch F., Hertig C., Rebmann G., Bull J., Dudler R.;
RT "A wheat glutathione-S-transferase gene with transposon-like
sequences in the promoter region.";
RT Plant Mol. Biol. 16:1089-1091(1991).
CC -!- FUNCTION: Conjugation of reduced glutathione to a wide number of
CC exogenous and endogenous hydrophobic electrophiles.
CC -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC -!- MISCELLANEOUS: GSTA2 is said to be a defective gene.
CC -!- SIMILARITY: Belongs to the GST superfamily. Phi family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP000006; BAA30738.1; ALT_INIT.
DR HAMAP; MF_00055; -; 1.
DR InterPro; IPR002737; DUF52.
DR Pfam; PF01875; UPF0103; 1.
DR ProDom; PD006364; DUF52; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 291 AA; 32448 MW; FB6FDFD19639005D CRC64;

Query Match 1.2%; Score 7; DB 1; Length 291;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 FKDLGEE 17
DB 26 FKDLGEE 32

RESULT 41
ANR2 MOUSE
ID ANR2 MOUSE STANDARD; PRT; 328 AA.
AC Q9WV06;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ankyrin repeat domain protein 2 (Skeletal muscle ankyrin repeat
DE protein) (mArpp).
DE protein.
GN ANKRD2 OR ARPP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Diaphragm;
RA Ievolella C., Formentin E., Lanfranchi G.;
RT "Characterization of a member of a new family proteins with ankyrin
RT repeats.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC STRAIN=129/Sv, and C57BL/10; TISSUE=Skeletal muscle, and Spleen;
RX MEDLINE=20334618; PubMed=10873377;
RA Kemp T.J., Sadusky T.J., Saltisi P., Carey N., Moss J., Yang S.Y.,
RA Sassoon D.A., Goldspink G., Coulton G.R.;
DE Hypothetical UPF0103 protein PH1626.
DE
```

RX stretch-responsive ankyrin-repeat protein." ;
 RA Genomics 66:229-241 (2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21999412; PubMed=12004005;
 RA Tsukamoto Y., Senda T., Nakano T., Nakada C., Hida T., Ishiguro N.,
 RA Kondo G., Baba T., Sato K., Oaki M., Mori S., Ito H., Moriyama M.,
 RA "Arpp, a new homolog of carp, is preferentially expressed in type 1
 RT skeletal muscle fibers and is markedly induced by denervation." ;
 RL Lab. Invest. 82:645-655 (2002).
 CC -!- FUNCTION: May play an important role in skeletal muscle
 CC hypertrophy.
 CC -!- TISSUE SPECIFICITY: Expressed in skeletal and cardiac muscles.
 CC -!- SIMILARITY: Contains 5 ANK repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AJ011118; CAB46646.1; -.
 CC EMBL; AJ249346; CAB99432.1; -.
 CC EMBL; AJ245514; CAB99431.1; -.
 CC HSP; P25963; 1IKN.
 CC MGD; MGI:1861447; Ankrd2.
 CC GO; GO:0030484; C:muscle fiber; ISS.
 CC GO; GO:0005634; C:nucleus; NAS.
 CC GO; GO:0030017; C:starcomere; IDA.
 CC GO; GO:0008307; F:structural constituent of muscle; ISS.
 CC GO; GO:0007517; F:muscle development; ISS.
 CC InterPro; IPR02110; ANK.
 CC Pfam; PF00023; ank; 4.
 CC PRINTS; PR01415; ANKYRIN.
 CC SMART; SM00248; ANK; 4.
 CC PROSITE; PS0088; ANK_REPEAT; 4.
 CC PROSITE; PS0297; ANK_REPEAT_REGION; 1.
 CC ANK repeat; Repeat.
 CC REPEAT 116 145 ANK 1.
 CC REPEAT 149 178 ANK 2.
 CC REPEAT 182 211 ANK 3.
 CC REPEAT 215 244 ANK 4.
 CC REPEAT 248 277 ANK 5.
 CC SEQUENCE 328 AA; 36707 MW; DB90D955EE9D175E CRC64;
 SQ
 Query Match 1.2%; Score 7; DB 1; Length 328;
 Best Local Similarity 100.0%; Pred.No. 63;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 130 NEETFLK 136
 DB 117 NEETFLK 123
 RESULT 42
 ID_QOR_MOUSE STANDARD; PRT; 331 AA.
 AC P47199; Q62508; Q99L63;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Quinone oxidoreductase (EC 1.6.5.5) (NADPH:quinone reductase) (zeta-
 DE crystallin).
 GN CRYZ.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 MEDLINE=94224126; PubMed=8170370;
 Gonzalez P., Hernandez-Galazadilla C., Rao P.V., Rodriguez I.R.,
 Ziegler J.S. Jr., Borras T.,
 "Comparative analysis of the zeta-crystallin/quinone reductase gene
 in guinea pig and mouse." ;
 Mol. Biol. Evol. 11:305-315 (1994).
 [2]
 SEQUENCE FROM N.A.
 TISSUE=Breast tumor;
 MEDLINE=22388257; PubMed=12477932;
 Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 Altschul S.F., Zengberg E., Buetow K.H., Smit A., Wang J., Hsieh F.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 Bosak S.A., McWay P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 Schnerch A., Schein J.E., Jones S.G.M., Marra M.A.,
 "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences." ;
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [3]
 SEQUENCE OF 130-331 FROM N.A.
 STRAIN=CS7BL/6J; TISSUE=Brain cortex;
 MEDLINE=96216731; PubMed=8645260;
 Kajiwara K., Nagasawa H., Shimizu-Nishikawa K., Ookura T., Kimura M.,
 Sugaya E.,
 "Molecular characterization of seizure-related genes isolated by
 differential screening." ;
 Biochem. Biophys. Res. Commun. 219:795-799 (1996).
 -!- FUNCTION: Does not have alcohol dehydrogenase activity. Binds NADP
 and acts through a one-electron transfer process. Orthoquinones
 are the best substrates. May act in the detoxification of
 xenobiotics (By similarity).
 -!- CATALYTIC ACTIVITY: NADPH + quinone = NADP(+) + semiquinone.
 -!- SUBUNIT: Homotetramer.
 -!- SURCELLULAR LOCATION: Cytoplasmic.
 -!- SIMILARITY: Belongs to the zinc-containing alcohol dehydrogenase
 family. Quinone oxidoreductase subfamily.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; S70056; AAB30620.2; -.
 CC EMBL; BC003800; AAH03800.1; -.
 CC EMBL; D78646; BAAL1463.1; -.
 CC ZIR; A54932; A54932.
 CC HSP; P28304; 1QOR.
 CC MGD; MGI:88527; Cryz.
 CC InterPro; IPR002085; Adh_zn_family.
 CC InterPro; IPR002364; QOR_zeta_crystal.
 CC Pfam; PF0107; Adh_zinc_N; 1.
 CC PROSITE; PS01162; QOR_ZETA_CRYSTAL; 1.
 CC Oxidoreductase; NADP; Zinc.
 CC CONFLICT 58 58 A -> T (IN REF. 2).
 CC CONFLICT 131 133 IPY -> TMD (IN REF. 3).
 CC SEQUENCE 331 AA; 35268 MW; 35816C043FE16A2 CRC64;
 SQ
 Query Match 1.2%; Score 7; DB 1; Length 331;

Best Local Similarity 100.0%; Pred. No. 63;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 570 EGGKLV 576
DB 183 EGGKLV 189

RESULT 43
CXA4 HUMAN
ID CXXA4 HUMAN STANDARD; PRT: 332 AA.
AC P35212; Q9P106; Q9UB11; Q9UN9A; Q9UNB0; Q9UNB1; Q9Y5N7;
DT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Gap junction alpha-4 protein (Connexin 37) (Cx37).
GN GJA4.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93195088; PubMed=7680674;
RA Reed K.B., Westphale E.M., Larson D.M., Wang H.-Z., Veenstra R.D.,
RA Beyer E.C.;
RT "Molecular cloning and functional expression of human connexin37, an
RT endothelial cell gap junction protein.";
RN J. Clin. Invest. 91:997-1004(1993).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS SER-70; VAL-127; ILE-129 AND SER-318.
RA van Zeijl L., Cotgreave I.A.;
RT "A connexin 37 genotypic variant in atherosclerosis.";
RN Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A., AND VARIANTS ILE-129 AND SER-318.
RA Kumari S., Varadraj K., Valiunas V., Ramanathan S.V., Beyer E.C.,
RA Brink P.R.;
RT "Functional expression and biophysical properties of two polymorphic
RT forms of human connexin37.";
RN Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Lench N.J., Williams G., Williams E., Gharani N., Franks S.;
RT "Connexin 37 mutation screening in anovulatory polycystic ovary
RT syndrome.";
RN Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Pearce A.;
RN Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas, and Spleen;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.T., Skalska U., Smalus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [7]
RP VARIANT ILE-129
RX MEDLINE=96358561; PubMed=8761439;
RA Krutovskikh V., Mironov N., Yamasaki H.;
RT "Human connexin 37 is polymorphic but not mutated in tumours.";
RN Carcinogenesis 17:1761-1763(1996).
RN [8]
RP VARIANT SER-318.
RX MEDLINE=99377098; PubMed=10447790;
RA Boerma M., Forsberg L., Van Zeijl L., Morgenstern R., De Faire U.,
RA Lemme C., Erlinge D., Thulin T., Hong Y., Cotgreave I.A.;
RT "A genetic polymorphism in connexin 37 as a prognostic marker for
RT atherosclerotic plaque development.";
RN J. Intern. Med. 246:211-218(1999).
RN [9]
RP VARIANT SER-318.
RX MEDLINE=20190940; PubMed=10728596;
RA Saito T., Krutovskikh V., Marion M.J., Ishak K.G., Bennett W.P.,
RA Yamasaki H.;
RT "Human hemangiosarcomas have a common polymorphism but no mutations in
RT the connexin37 gene.";
RN Int. J. Cancer 86:67-70(2000).
CC -!- FUNCTION: One gap junction consists of a cluster of closely packed
CC pairs of transmembrane channels, the connexons, through which
CC materials of low MW diffuse from one cell to a neighboring cell.
CC -!- SUBUNIT: A connexon is composed of a hexamer of connexins.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Expressed in multiple organs and tissues,
CC including heart, uterus, ovary, and blood vessel endothelium.
CC -!- SIMILARITY: Belongs to the connexin family. Alpha-type (group II)
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M96789; AAAS2558.2; -;
CC EMBL; AF139100; AAD31869.1; -;
CC EMBL; AF139101; AAD31870.1; -;
CC EMBL; AF139102; AAD31871.1; -;
CC EMBL; AF139103; AAD31872.1; -;
CC EMBL; AF139104; AAD31873.1; -;
CC EMBL; AF139105; AAD31874.1; -;
CC EMBL; AF181620; AAD56940.1; -;
CC EMBL; AF132674; AAF62342.1; -;
CC EMBL; AL121988; CAB90268.1; -;
CC EMBL; BC027889; AB27889.1; -;
CC Genew; HGNC:4278; GJA4.
CC MIN; 121012; -;
CC GO; GO:0005921; C:gap junction; TAS.
CC GO; GO:0005887; C:integral to plasma membrane; TAS.
CC GO; GO:0015285; F:connexon channel activity; TAS.
CC GO; GO:0007043; F:intercellular junction assembly; TAS.
CC GO; GO:0006810; P:transport; TAS.
CC InterPro; IPR000500; Connexin.
CC Pfam; PF00029; connexin; 1.
CC PRINTS; PR00206; CONNEXIN.
CC SMART; SM00037; CNX; 1.
CC PROSITE; PS00407; CONNEXINS_1; 1.
CC PROSITE; PS00408; CONNEXINS_2; 1.
CC Gap junction; Transmembrane; Polymorphism.
CC INT_MET 0 0 BY SIMILARITY.
FT DOMAIN 1 19 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 20 39 POTENTIAL.
FT DOMAIN 40 75 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 76 98 POTENTIAL.
FT DOMAIN 99 147 CYTOPLASMIC (POTENTIAL).

HAMAP: MF 00016; -, 1.
 InterPro: IPR003593; AAA_ATPase.
 InterPro: IPR003959; AAA_ATPase_centre.
 InterPro: IPR004605; RuvB.
 Pfam: PF00004; AAA; 1.
 SMART: SM00382; AAA; 1.
 TIGRFAMs: TIGR00635; ruvB; 1.
 DNA repair; SOS response; ATP-binding; DNA recombination; Helicase;
 Complete proteome.
 NP_BIND 57 64 ATP (POTENTIAL).
 SEQUENCE 337 AA; 37331 MW; C39DB1FE0B318455 CRC64;

Query March 1.2%; Score 7; DB 1; Length 337;
 Best Local Similarity 100.0%; Pred.No. 64; 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 LYEIARR 145
 |||||
 DB 204 LYEIARR 210

RESULT 45
 ADD_STRVG ADD_STRVG STANDARD; PRT; 339 AA.
 ID ADD_STRVG
 AC P53984;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Adenosine deaminase (EC 3.5.4.4) (Adenosine aminohydrolase).
 GN ADD OR ADA.
 OS Streptomyces virginiae.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 NCBI_TaxID=1961;
 [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=96257210; PubMed=8675024;
 RX Katayama M., Sakai Y., Okamoto S., Ihara F., Nihira T., Yamada Y.;
 RT "Gene organization in the ada-rplL region of Streptomyces virginiae.";
 RL Gene 171:135-136(1996).
 CC -!- CATALYTIC ACTIVITY: Adenosine + H(2)O = inosine + NH(3).
 CC -!- SIMILARITY: Belongs to the adenosine and AMP deaminases family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: D50624; BAA0298.1; -.
 CC PIR: T11785; T11785.
 DR HAMAP: MF 00540; -, 1.
 DR InterPro: IPR006650; A/AMP deam AS.
 DR InterPro: IPR001365; A/AMP_deaminase.
 DR InterPro: IPR006330; A_deaminase.
 DR Pfam: PF00962; A_deaminase; 1.
 DR TIGRFAMs: TIGR01430; aden deam; 1.
 DR PROSITE: PS00485; A_DEAMINASE; 1.
 DR Hydroxylase; Nucleotide metabolism.
 FT ACT_SITE 200 300 POTENTIAL.
 FT ACT_SITE 248 248 POTENTIAL.
 FT ACT_SITE 281 281 POTENTIAL.
 FT ACT_SITE 282 282 POTENTIAL.
 FT SEQUENCE 339 AA; 37181 MW; 7C1C221FB927E5AD CRC64;

Query March 1.2%; Score 7; DB 1; Length 339;
 Best Local Similarity 100.0%; Pred.No. 65;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 YEIARRH 146
 |||||

```
Db      294 YEIARRH 300
RESULT 46
RECA_ENTFA
AC      P42444; STANDARD; PRT; 348 AA.
DT      01-NOV-1995 (Rel. 32, Created)
DT      10-OCT-2003 (Rel. 42, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      ReCA protein (Recombinase A).
GN      RECA OR EF3171.
OS      Enterococcus faecalis (Streptococcus faecalis).
OC      Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX      NCBI_TaxID=1351;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN=V583 / ATCC 700802;
RX      MEDLINE=22550857; PubMed=12663927;
RA      Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA      Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA      Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
RA      Daugherty S., DeBoy R.T., Kolonay J., Madupu R., Khouri H.,
RA      Vanathavan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,
RA      Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;
RA      "Role of mobile DNA in the evolution of vancomycin-resistant
RT      Enterococcus faecalis.";
RL      Science 299:2071-2074(2003).
RN      [2]
RP      SEQUENCE OF 91-192 FROM N.A.
RX      MEDLINE=92210521; PubMed=1556091;
RA      Dybvig K., Hollingshead S.K., Heath D.G., Clewell D.B., Sun F.,
RA      Woodard A.;
RT      "Degenerate oligonucleotide primers for enzymatic amplification of
RT      reCA sequences from Gram-positive bacteria and mycoplasmas.";
RL      J. Bacteriol. 174:2729-2732(1992)
CC      -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC      single-stranded DNA, the ATP-dependent uptake of single-stranded
CC      DNA by duplex DNA, and the ATP-dependent hybridization of
CC      homologous single-stranded DNAs. It interacts with lexA causing
CC      its activation and leading to its autocatalytic cleavage.
CC      -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC      -!- SIMILARITY: Belongs to the reCA family.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      EMBL; AE016957; AA082845.1; -.
CC      DR      EMBL; M81466; AAA24782.1; -.
CC      DR      PIR; C41846; C41846.
CC      DR      HSSP; P26345; 1G18.
CC      DR      TIGR; EF3171; -.
CC      DR      HAVAP; WP_002688; -.
CC      DR      InterPro; IPR003593; AAA ATPase.
CC      DR      InterPro; IPR001553; RecA.
CC      DR      Pfam; PF00154; reCA; 1.
CC      DR      PRINTS; PR00142; RECA.
CC      DR      ProDom; PD000229; RecA; 1.
CC      DR      SMART; SM00382; AAA; 1.
CC      DR      PROSITE; PS00321; RECA_1; 1.
CC      DR      PROSITE; PS50162; RECA_2; 1.
CC      DR      PROSITE; PS50163; RECA_3; 1.
CC      DR      DNA damage; DNA recombination; SOS response; ATP-binding; DNA-binding;
KW      Complete proteome.
FT      NP_BIND 65 72 ATP (BY SIMILARITY).
SQ      SEQUENCE 348 AA; 37409 NW; E8D4F5ED406842F6 CRC64;
Query Match 1.2%; Score 7; DB 1; Length 348;
```

```
Best Local Similarity 100.0%; Pred. No. 66;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      229 AEVSKLV 235
DB      316 AEVSKLV 322
|||||
RESULT 47
CD68_HUMAN
ID      CD68_HUMAN STANDARD; PRT; 354 AA.
AC      P34810; Q96BI7;
DT      01-FEB-1994 (Rel. 28, Created)
DT      01-FEB-1994 (Rel. 28, Last sequence update)
DT      15-MAR-2004 (Rel. 43, Last annotation update)
DE      Macrophage precursor (CD68 antigen) (GP110).
GN      CD68.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX      MEDLINE=93200523; PubMed=7680921;
RA      Holmes C.L., Simmons D.L.;
RT      "Molecular cloning of CD68, a human macrophage marker related to
RT      lysosomal glycoproteins.";
RL      Blood 81:1607-1613(1993).
RN      [2]
RP      SEQUENCE FROM N.A., AND VARIANT GLN-254.
RX      TISSUE=Skin;
RX      MEDLINE=22388257; PubMed=12477932;
RA      Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan K., Moore T., Max S.I., Wang J., Hsieh P.,
RA      Diatchenko L., Marudina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Tsien T.B., Toshiyuki S., Carninci P., Prange C.,
RA      Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA      Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA      Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA      Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA      Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length
RT      human and mouse cDNA sequences.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN      [3]
RP      SEQUENCE OF 1-28 FROM N.A.
RX      MEDLINE=99009345; PubMed=9790779;
RA      Jones E., Quinn C.M., See C.G., Montgomery D.S., Ford M.J.,
RA      Koelble K., Gordon S., Greaves D.R.;
RT      "The linked human elongation initiation factor 4A1 (EIF4A1) and CD68
RL      genes map to chromosome 17p13.";
CC      Genomics 53:248-250(1998).
CC      -!- FUNCTION: Could play a role in phagocytic activities of tissue
CC      macrophages, both in intracellular lysosomal metabolism and
CC      extracellular cell-cell and cell-pathogen interactions. Bind to
CC      tissue- and organ-specific lectins or selectins, allowing homing
CC      of macrophage subsets to particular sites. Rapid recirculation of
CC      CD68 from endosomes, lysosomes to the plasma membrane may allow
CC      macrophages to crawl over selectin bearing substrates or other
CC      cells.
CC      -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ENDOSOMAL OR
CC      LYSOSOMAL (LONG VARIANT) AND TO A LESSER EXTENT ON THE CELL
CC      SURFACE (SHORT VARIANT).
CC      -!- ALTERNATIVE PRODUCTS:
CC      Event=Alternative splicing; Named isoforms=2;
```

```
CC Name=Long;
CC IsoId=P34810-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=P34810-2; Sequence=VSP_003041, VSP_003042;
CC TISSUE SPECIFICITY: Highly expressed by blood monocytes and tissue
CC macrophages. Also expressed in many tumor cell lines which could
CC allow them to attach to selections on vascular endothelium,
CC facilitating their dissemination to secondary sites.
CC PTM: N- and O-glycosylated.
CC SIMILARITY: Belongs to the LAMP family.
CC DATABASE: NAME=PROW; NOTE=CD guide CD68 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd68.htm".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; S57235; AAB25811.1; -
CC GO; GO:0016021; C:integral to membrane; TAS.
CC EMBL; BC015557; AAH15557.1; -
CC EMBL; AF060540; AAC70006.1; -
CC FIC; A48931; A48931
CC Genew; HGNC:1693; CD68.
CC MIM; 153634; -
CC GO; GO:0016021; C:integral to membrane; TAS.
CC GO; GO:0005624; C:membrane fraction; TAS.
CC InterPro; IPR002000; Lamp.
CC Pfam; PF01299; Lamp; 1.
CC PRINTS; PR00336; LYSGASOCTDMP.
CC PROSITE; PS00311; LAMP_2; 1.
CC Transmembrane; Glycoprotein; Signal; Lysosome; Repeat; Antigen;
KW Alternative splicing; Polymorphism.
FT SIGNAL 1 21
FT CHAIN 22 354
FT MACROPHAGALIN.
FT DOMAIN 22 319
FT TRANSMEM 320 344
FT DOMAIN 345 354
FT DOMAIN 23 140
FT DOMAIN 140 152
FT DOMAIN 70 129
FT REPEAT 70 99
FT REPEAT 100 129
FT REPEAT 169 207
FT DISULFID 277 314
FT CARBOHYD 88 88
FT CARBOHYD 96 96
FT CARBOHYD 118 118
FT CARBOHYD 126 126
FT CARBOHYD 164 164
FT CARBOHYD 199 199
FT CARBOHYD 246 246
FT CARBOHYD 261 261
FT CARBOHYD 279 279
FT VARSPLIC 17 43
FT VARSPLIC 83 112
FT VARIAT 254 254
FT VARIAT 340 340
FT SEQUENCE 354 AA; 37408 MW; 0A29ACBBF9431B0F CRC64;
SQ
Query Match 1.2%; Score 7; DB 1; Length 354;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 ALVLIAF 27
DB 335 ALVLIAF 341
```

RESULT 48

```
OMPF_SALTI
ID OMPF_SALTI STANDARD; PRT; 363 AA.
AC Q56113;
DT 15-DEC-1998 (Rel. 37, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Outer membrane protein F precursor (Porin ompF) (Outer membrane
DE protein S3).
GN OMPF OR OMP83 OR STY1002 OR T1935.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IMSS-1;
RA Fernandez-Mora M., Calva E.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.F., Holroyd S., Jagsis K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18."
RL Nature 413:848-852(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodyanni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18."
RL J. Bacteriol. 185:2330-2337(2003).
CC -!- FUNCTION: ompF is a porin that forms passive diffusion pores which
CC allow small molecular weight hydrophilic materials across the
CC outer membrane. It is also a receptor for the bacteriophage T2 (By
CC similarity).
CC -!- SUBUNIT: Homotrimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC -!- SIMILARITY: Belongs to the Gram-negative porin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X89757; CAA61905.1; -
CC EMBL; AL627268; CAD05399.1; -
CC EMBL; AB016840; AAO59550.1; -
CC HSSP; P02931; IGFN.
CC InterPro; IPR003229; OMP 2.
CC InterPro; IPR001702; Porin Gram-ve.
CC Pfam; PF00267; Gram-ve porins; 1.
CC PRINTS; PR00182; ECOLNEIPORIN.
CC ProDom; PD000808; OMP 2; 1.
CC PROSITE; PS00576; GRAM_NEG_PORIN; 1.
KW Outer membrane; Transmembrane; Porin; Phage recognition; Signal;
KW Complete proteome.
```

```

FT SIGNAL 1 22 BY SIMILARITY.
FT CHAIN 23 363 OUTER MEMBRANE PROTEIN F.
FT TRANSMEM 29 28 BY SIMILARITY.
FT DOMAIN 29 28 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 30 45 BY SIMILARITY.
FT DOMAIN 46 56 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 57 69 BY SIMILARITY.
FT DOMAIN 70 71 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 72 84 BY SIMILARITY.
FT DOMAIN 85 98 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 100 108 BY SIMILARITY.
FT DOMAIN 109 109 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 110 117 BY SIMILARITY.
FT DOMAIN 118 154 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 155 161 BY SIMILARITY.
FT DOMAIN 162 169 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 170 181 BY SIMILARITY.
FT DOMAIN 182 192 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 193 203 BY SIMILARITY.
FT DOMAIN 204 204 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 205 217 BY SIMILARITY.
FT DOMAIN 218 230 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 231 242 BY SIMILARITY.
FT DOMAIN 243 243 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 244 256 BY SIMILARITY.
FT DOMAIN 257 272 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 273 285 BY SIMILARITY.
FT DOMAIN 286 287 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 288 301 BY SIMILARITY.
FT DOMAIN 302 312 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 313 324 BY SIMILARITY.
FT DOMAIN 325 326 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 327 336 BY SIMILARITY.
FT DOMAIN 337 353 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 354 363 BY SIMILARITY.
FT CONFLICT 307 307 D -> G (IN REF. 1).
SQ SEQUENCE 363 AA; 40106 MW; F5059B37EAS16859 CRC64;

Query Match 1.2%; Score 7; DB 1; Length 363;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 258 ADLAKYI 264
DB 310 ADLAKYI 316
|||||

RESULT 49
ID OMFP_SALTY STANDARD; PRT; 363 AA.
AC P37432;
DT 01-OCT-1994 (Rel. 30, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Outer membrane protein F precursor (Porin ompF) (Outer membrane
DE protein 1A) (Outer membrane protein 1A) (Outer membrane protein B).
GN OMFP OR STM0999.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
CX NCBI_TaxID=602;
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=lr2;
RA Venegas A.; Gomez I.; Bruce E.; Martinez M.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M.; Sanderson K.E.; Spieth J.; Clifton S.W.; Latreille P.;
RA Courtney L.; Porwollik S.; Ali J.; Dante M.; Du F.; Hou S.; Layman D.;
RA Leonard S.; Nguyen C.; Scott K.; Holmes A.; Grewal N.; Mulvaney E.;

```

```

RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856(2001).
CC -!- FUNCTION: OMFP IS A PORIN THAT FORMS PASSIVE DIFFUSION PORES WHICH
CC ALLOW SMALL MOLECULAR WEIGHT HYDROPHILIC MATERIALS ACROSS THE
CC OUTER MEMBRANE. IT IS ALSO RECEPTOR FOR THE BACTERIOPHAGE T2 (BY
CC similarity).
CC -!- SUBUNIT: Homotrimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane
CC (By similarity).
CC -!- SIMILARITY: Belongs to the Gram-negative porin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z31594; CA83471.1; -.
CC EMBL; AB008743; AAL19933.1; -.
CC PIR; S43159; S43159.
CC HSSP; P02931; IGFN.
CC StyGene; SG10264; ompF.
CC InterPro; IPR003229; OMP_2.
CC InterPro; IPR001702; Porin Gram-ve.
CC Pfam; PF00267; Gram-ve porins; 1.
CC PRINTS; PR00182; SCOLNEIPORIN.
CC ProDom; PD000808; OMP_2; 1.
CC PROSITE; PS00576; GRAM NEG PORIN; 1.
CC Outer membrane; Transmembrane; Porin; Phage recognition; Signal;
CC Complete proteome.
FT SIGNAL 1 22 BY SIMILARITY.
FT CHAIN 23 363 OUTER MEMBRANE PROTEIN F.
FT TRANSMEM 29 28 BY SIMILARITY.
FT DOMAIN 29 28 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 30 45 BY SIMILARITY.
FT DOMAIN 46 56 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 57 69 BY SIMILARITY.
FT DOMAIN 70 71 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 72 84 BY SIMILARITY.
FT DOMAIN 85 99 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 100 108 BY SIMILARITY.
FT DOMAIN 109 109 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 110 117 BY SIMILARITY.
FT DOMAIN 118 154 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 155 161 BY SIMILARITY.
FT DOMAIN 162 169 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 170 181 BY SIMILARITY.
FT DOMAIN 182 192 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 193 203 BY SIMILARITY.
FT DOMAIN 204 204 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 205 217 BY SIMILARITY.
FT DOMAIN 218 230 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 231 242 BY SIMILARITY.
FT DOMAIN 243 243 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 244 256 BY SIMILARITY.
FT DOMAIN 257 272 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 273 285 BY SIMILARITY.
FT DOMAIN 286 287 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 288 301 BY SIMILARITY.
FT DOMAIN 302 312 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 313 324 BY SIMILARITY.
FT DOMAIN 325 326 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 327 336 BY SIMILARITY.
FT DOMAIN 337 353 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 354 363 BY SIMILARITY.
FT CONFLICT 232 232 E -> A (IN REF. 1).
FT CONFLICT 343 343 T -> V (IN REF. 1).
FT CONFLICT 360 360
SQ SEQUENCE 363 AA; 40048 MW; F6069B34E9516859 CRC64;

```

Query Match 1.2%; Score 7; DB 1; Length 363;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 258 ADLAKYI 264
|||||
DB 310 ADLAKYI 316

RESULT 50
AROB_CANBF STANDARD; PRT; 364 AA.
AC Q7VRN3;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 3-dehydroquininate synthase (EC 4.2.3.4).
GN AROB OR BFL571.

OS Candidatus Blochmannia floridanus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; ant endosymbionts; Candidatus Blochmannia.
OX NCBI_TaxID=203907;
[1]_TaxID=203907;
RN SEQUENCE FROM N.A.
RP MEDLINE=22784745; PubMed=12886019;
RX Gil R., Silva F.J., Zientz E., Delmotte F., Gonzalez-Candelas F.,
RA Latorre A., Rausell C., Kamerbeek J., Gadau J., Hoelldobler B.,
RA van Ham R.C.H.J., Gross R., Moya A.;
RT "The genome sequence of Blochmannia floridanus: comparative analysis
of reduced genomes";
RL Proc Natl Acad Sci U S A. 100:9388-9393(2003).
CC -1- CATALYTIC ACTIVITY: 3-deoxy-arabino-heptulosonate 7-phosphate = 3-
CC dehydroquininate + phosphate.
CC -1- COFACTOR: NAD and a divalent metal cation (By similarity).
CC -1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC second step.
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: Belongs to the dehydroquininate synthase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; BX248586; CAD83253.1; -;
CC HAMAP; MF 00110; -; 1.
CC InterPro; IPR002658; DHQ synthase.
CC Pfam; PF01761; DHQ synthase; 1.
CC TIGRFAMs; TIGR01357; arob; 1.
CC KW Aromatic amino acid biosynthesis; Lyase; NAD; Complete proteome.
CC SEQUENCE 364 AA; 40644 MW; 30C0191A2509BB75 CRC64;

Query Match 1.2%; Score 7; DB 1; Length 364;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 514 CTUSEKE 520
|||||
DB 173 CTUSEKE 179

RESULT 51
PYRC_ARATH STANDARD; PRT; 377 AA.
AC Q04904;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Dihydroorotase, mitochondrial precursor (EC 3.5.2.3) (DHOase).
GN PYR4 OR AT4G22930 OR F7H19.110.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN SEQUENCE FROM N.A.
RP STRAIN=cv. Columbia;
RC Zhou L., Lacroute F., Thornburg R.W.;
RA "Characterization of the Arabidopsis thaliana cDNA encoding
RT dihydroorotase";
RL (In) Plant Gene Register PGR97-115.
RN SEQUENCE FROM N.A.
RP STRAIN=cv. Columbia;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger W.,
RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hebeisel J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McQuillagh B., Bilham L., Robben J., Vandenbussche F.,
RA Van der Schueren J., Grynoprez B., Chuang Y.-J., Vandenbussche F.,
RA Braeken M., Welljens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA Weitenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA Holzer E., Brandt A., Peters S., van Staveren M., Dirke W.,
RA Moeljan P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McLeay K., Mayes R.,
RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
RA Dose S., de Haan M., Maars A.C., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fattmann B., Grandjean K., Dauner D., Herzl A.,
RA Neumann S., Argitoni A., Vitale D., Liguori R., Piravandi E.,
RA Vassennat O., Quigley F., Clabaud G., Muendlein A., Aubourg S.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Casacuberta E.,
RA Gibbons T., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Frances P., Bielke C.,
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Farnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threiden J.,
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.W.,
RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shan R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA Grant S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA Chen E., Marra M.A., Martienssen R., McCombie W.R.;
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT thaliana";
RL Nature 402:769-777(1999).
CC -1- CATALYTIC ACTIVITY: (S)-dihydroorotate + H(2)O = N-carbamoyl-L-
CC aspartate.
CC -1- COFACTOR: Binds 2 zinc ions per subunit (By similarity).
CC -1- PATHWAY: Pyrimidine biosynthesis; third step.
CC -1- SUBCELLULAR LOCATION: Mitochondrial (Potential).
CC -1- SIMILARITY: Belongs to the DHOase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; BX248586; CAD83253.1; -;
CC HAMAP; MF 00110; -; 1.
CC InterPro; IPR002658; DHQ synthase.
CC Pfam; PF01761; DHQ synthase; 1.
CC TIGRFAMs; TIGR01357; arob; 1.
CC KW Aromatic amino acid biosynthesis; Lyase; NAD; Complete proteome.
CC SEQUENCE 364 AA; 40644 MW; 30C0191A2509BB75 CRC64;

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch)

DR EMBL; AF000146; BAB71134.1; -;
DR EMBL; AL031018; CAB19808.1; -;
DR EMBL; AL161598; CAB79248.1; -;
DR PIR; T05124; T05124.
DR InterPro; IPR006680; Amidohydro 1.
DR InterPro; IPR004721; Pept_M38_DHodimr.
DR InterPro; IPR002195; Pept_M38_nph.
DR Pfam; PF01979; Amidohydro_1; 1.
DR TIGSFams; TIGR00856; PyrcDimer; 1.
DR PROSITE; PS00482; DIHYDROOROTASE_1; 1.
DR PROSITE; PS00483; DIHYDROOROTASE_2; 1.
DR Pyrimidine biosynthesis; Hydrolase; Metal-binding; Zinc;
KW Transit peptide; Mitochondrion.
FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
FT CHAIN ? 377 DIHYDROOROTASE.
FT METAL 44 44 ZINC 1 (BY SIMILARITY).
FT METAL 46 46 ZINC 1 (BY SIMILARITY).
FT METAL 130 130 ZINC 1 AND 2 (BY SIMILARITY).
FT METAL 168 168 ZINC 2 (BY SIMILARITY).
FT METAL 206 206 ZINC 2 (BY SIMILARITY).
FT METAL 280 280 ZINC 1 (BY SIMILARITY).
SQ SEQUENCE 377 AA; 41849 MW; 783C22B85581DEB8 CRC64;

Query Match 1.2%; Score 7; DB 1; Length 377;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 370 YAKVFDE 376
|||||
DB 308 YAKVFDE 314

RESULT 52
CEGT_HUMAN
ID CEGT_HUMAN STANDARD; PRT; 394 AA.
AC Q16739.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ceramide glucosyltransferase (EC 2.4.1.80) (Glucosylceramide synthase)
DE (GCS) (UDP-glucose:N-acylsphingosine D-glucosyltransferase) (UDP-
DE glucose ceramide glucosyltransferase) (GLCT-1).
GN UGCG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96209784; PubMed=8643456;
RA Ichikawa S., Sakiyama H., Suzuki G., Hidari K.I.-P., Hirabayashi Y.;
RT "Expression cloning of a cDNA for human ceramide glucosyltransferase
RT that catalyzes the first glycosylation step of glycosphingolipid
RT synthesis.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:4638-4643 (1996).
RN [2]
RN ERRATUM.
RX MEDLINE=97057299; PubMed=8901638;
RA Ichikawa S., Sakiyama H., Suzuki G., Hidari K.I.-P., Hirabayashi Y.;
RL Proc. Natl. Acad. Sci. U.S.A. 93:12654-12654 (1996).
RN [3]
RN SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: May serve as a "flippase" as well as a
CC Glucosyltransferase that transfers glucose to ceramide.
CC -!- CATALYTIC ACTIVITY: UDP-glucose + N-acylsphingosine = UDP + D-
CC glucosyl-N-acylsphingosine.
CC -!- PATHWAY: Glycosphingolipid synthesis; first glycosylation step.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum.
CC -!- TISSUE SPECIFICITY: Found in all tissues examined.
CC -!- SIMILARITY: Belongs to the glucosyltransferase family 2.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch)

DR EMBL; D50840; BAA09451.1; -;
DR EMBL; BC038711; AAH38711.1; -;
DR Genew; HGNC:12524; UGCG.
DR MIN; 602874; -;
DR GO; GO:0005624; C:membrane fraction; TAS.
DR GO; GO:0008120; P:ceramide glucosyltransferase activity; TAS.
DR GO; GO:0008544; P:epidermal differentiation; TAS.
DR GO; GO:0006679; P:glucosylceramide biosynthesis; TAS.
DR GO; GO:0006688; P:glycosphingolipid biosynthesis; TAS.
DR InterPro; IPR001173; Glyco_transf_2;
DR Pfam; PF00535; Glycos_transf_2; 1.
KW Transferase; Glucosyltransferase; Transmembrane; Signal-anchor;
KW Endoplasmic reticulum.
FT DOMAIN 1 10 LUMENAL (POTENTIAL).
FT TRANSMEM 11 31 SIGNAL-ANCHOR (POTENTIAL).
FT TRANSMEM 286 306 POTENTIAL.
FT TRANSMEM 314 334 POTENTIAL.
SQ SEQUENCE 394 AA; 44853 MW; 3B998569F8A96449 CRC64;

Query Match 1.2%; Score 7; DB 1; Length 394;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 LIAPAQY 30
|||||
DB 226 LIAPAQY 232

RESULT 53
CEGT_MOUSE
ID CEGT_MOUSE STANDARD; PRT; 394 AA.
AC O88693;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ceramide glucosyltransferase (EC 2.4.1.80) (Glucosylceramide synthase)
DE (GCS) (UDP-glucose:N-acylsphingosine D-glucosyltransferase) (UDP-
DE glucose ceramide glucosyltransferase) (GLCT-1).

GN UGCG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RA MEDLINE=98285232; PubMed=9623774;
RX Ichikawa S., Ozawa K., Hirabayashi Y.;
RA "Molecular cloning and expression of mouse ceramide
glucosyltransferase";
RT Biochem. Mol. Biol. Int. 44:1193-1202(1998).
RL [2]
RN SEQUENCE FROM N.A.
RC STRAIN=129/SvJ;
RA MEDLINE=99119072; PubMed=9918791;
RX Ichikawa S., Ozawa K., Hirabayashi Y.;
RA "Molecular cloning and characterization of the mouse ceramide
glucosyltransferase gene";
RT Biochem. Biophys. Res. Commun. 253:707-711(1998).
RL [3]
RN SEQUENCE FROM N.A.
RC TISSUE=Limb;
RA MEDLINE=22388257; PubMed=12477932;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lomellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A.C., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,
RA Scherth A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP FUNCTION: May serve as a "flippase" as well as a
glucosyltransferase that transfers glucose to ceramide.
CC -!- CATALYTIC ACTIVITY: UDP-glucose + N-acylsphingosine = UDP + D-
glucosyl-N-acylsphingosine.
CC -!- PATHWAY: Glycosphingolipid synthesis; first glycosylation
step.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
reticulum (By similarity).
CC -!- SIMILARITY: Belongs to the glucosyltransferase family 2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC -----
CC ENBL; D9866; BAA28792.1; --
CC ENBL; AB012807; BAA33558.1; JOINED.
CC ENBL; AB012799; BAA33558.1; JOINED.
CC ENBL; AB012800; BAA33558.1; JOINED.
CC ENBL; AB012801; BAA33558.1; JOINED.
CC ENBL; AB012802; BAA33558.1; JOINED.
CC ENBL; AB012803; BAA33558.1; JOINED.
CC ENBL; AB012804; BAA33558.1; JOINED.
CC ENBL; AB012805; BAA33558.1; JOINED.
CC ENBL; AB012806; BAA33558.1; JOINED.

DR ENBL; BC050828; AAH50828.1; --
DR MGD; MGI:1332243; UGCG.
DR InterPro; IPR001173; Glyco_transf_2.
DR Pfam; PF00535; Glycosyltransferase; Transmembrane;
KW Signal-anchor; Endoplasmic reticulum.
KW DOMAIN 1 10 LUMENAL (POTENTIAL).
FT TRANSMEM 11 31 POTENTIAL.
FT TRANSMEM 287 307 POTENTIAL.
FT TRANSMEM 314 334 POTENTIAL.
SQ SEQUENCE 394 AA; 44838 MW; 8B18D09437CACE0E CRC64;
Query Match 1.2%; Score 7; DB 1; Length 394;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 24 LIAPAQY 30
DB 226 LIAPAQY 232
RESULT 54
CEGT_RAT ID CEGT_RAT STANDARD; PRT; 394 AA.
AC Q9R0E0; O55149;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ceramide glucosyltransferase (EC 2.4.1.80) (Glucosylceramide synthase)
DE (GCS) (UDP-glucose:N-acylsphingosine D-glucosyltransferase) (UDP-
glucose ceramide glucosyltransferase) (GLCT-1).
GN UGCG.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.; AND FUNCTION.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=99321486; PubMed=10393098;
RA Wu K., Marks D.L., Watanabe R., Paul P., Rajan N., Pagano R.E.;
RT "Histidine-193 of rat glucosylceramide synthase resides in a UDP-
glucose- and inhibitor (D-threo-1-phenyl-2-decanoylamino-3-
morpholinopropan-1-ol)-binding region: a biochemical and mutational
study.";
RL Biochem. J. 341:395-400(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Liver;
RA Bunning C., Orci L., Hirabayashi Y., Wieland F.T., Jeckel D.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May serve as a "flippase" as well as a
glucosyltransferase that transfers glucose to ceramide. Able to
use UDP-galactose to synthesize galactosylceramide with 10% of
efficiency with which it utilizes UDP-glucose.
CC -!- CATALYTIC ACTIVITY: UDP-glucose + N-acylsphingosine = UDP + D-
glucosyl-N-acylsphingosine.
CC -!- PATHWAY: Glycosphingolipid synthesis; first glycosylation step.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
reticulum.
CC -!- SIMILARITY: Belongs to the glucosyltransferase family 2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC -----
CC ENBL; AF047707; AAD02464.1; --
CC ENBL; AJ224156; CAA11853.1; --
DR InterPro; IPR001173; Glyco_transf_2.

DR Pfam: PF00535; Glycosyltransferase; Transmembrane; Signal-anchor;
KW Transferase; Glycosyltransferase; Polymorphism.
FT DOMAIN 1 10 LUMENAL (POTENTIAL).
FT TRANSMEM 11 31 SIGNAL-ANCHOR (POTENTIAL).
FT TRANSMEM 286 306 POTENTIAL.
FT TRANSMEM 314 334 POTENTIAL.
FT VARIANT 8 8 Q -> L (IN STRAIN WISTAR).
FT VARIANT 89 89 D -> G (IN STRAIN WISTAR).
FT VARIANT 153 153 T -> S (IN STRAIN WISTAR).
FT VARIANT 179 179 G -> A (IN STRAIN WISTAR).
FT VARIANT 387 387 T -> I (IN STRAIN WISTAR).
SQ SEQUENCE 394 AA; 44822 MW; 214581C0B8D9152C CRC64;

Query Match 1.2%; Score 7; DB 1; Length 394;
Best Local Similarity 100.0%; Pred. No. 73; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;

QY 24 LIAPAQY 30
DB 226 LIAPAQY 232

RESULT 55
PQOE BRAJA STANDARD; PRT; 399 AA.
AC Q89FG1;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Coenzyme PQQ synthesis protein E (Pyrroloquinoline quinone biosynthesis protein E).
DE PQOE OR BLR6739.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T., Sasamoto S., Watanabe A., Ideawa K., Iriiguchi M., Kawashina K., Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M., Tabata S.;
RA "Complete genomic sequence of nitrogen-fixing symbiotic bacterium Bradyrhizobium japonicum USDA110."
RL DNA Res. 9:189-197(2002).
CC -|- COPACTOR: Iron-sulfur cluster (Potential).
CC -|- PATHWAY: Pyrroloquinoline quinone (PQQ) biosynthesis.
CC -|- SIMILARITY: Belongs to the radical SAM superfamily. PqqE family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/seqdb/doc/seqdb.html> or send an email to license@ebi.ac.uk).
CC
CC EMBL: AF005959; BACS2004.1; --
DR HAMAP: MF 00660; -- 1.
DR InterPro: IPR000385; MoA NifB PqqE.
DR InterPro: IPR007197; Radical SAM.
DR Pfam: PF04055; Radical SAM; 1
DR PROSITE: PS01305; MOA_NIFB_PQOE; 1.
KW PQQ biosynthesis; Iron-sulfur; Complete proteome.
FT METAL 45 45 IRON-SULFUR (POTENTIAL).
FT METAL 49 49 IRON-SULFUR (POTENTIAL).
FT METAL 52 52 IRON-SULFUR (POTENTIAL).
SQ SEQUENCE 399 AA; 44607 MW; 09694B8A97971451 CRC64;

Query Match 1.2%; Score 7; DB 1; Length 399;

Best Local Similarity 100.0%; Pred. No. 74; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;

QY 529 LVELVKH 535
DB 98 LVELVKH 104

RESULT 56
RENS_MOUSE STANDARD; PRT; 401 AA.
ID RENS_MOUSE
AC P00796; F70229; P97955; Q62155;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Renin 2 precursor (EC 3.4.23.15) (Angiotensinogenase) (Submandibular gland rein).
DE Gland rein.
GN REN2 OR REN-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 64-351 AND 354-401.
RX MEDLINE=83014991; PubMed=6812055;
RA Misono K.S., Chang J.-J., Inagami T.;
RT "Amino acid sequence of mouse submaxillary gland renin."
RL Proc. Natl. Acad. Sci. U.S.A. 79:4858-4862(1982).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=82220074; PubMed=6283373;
RA Panthier J.-J., Foote S., Chambraud B., Strosberg A.D., Corvol P., Rougeon F.;
RT "Complete amino acid sequence and maturation of the mouse submaxillary gland renin precursor."
RL Nature 298:90-92(1982).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=FVS/N; TISSUE=Salivary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullen S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Wooley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 1-29 FROM N.A.
RX MEDLINE=84298161; PubMed=6089205;
RA Panthier J.-J., Dreyfus M., Roux D.T.L., Rougeon F.;
RT "Mouse kidney and submaxillary gland renin genes differ in their 5' putative regulatory sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 81:5489-5493(1984).
RN [5]
RP SEQUENCE OF 1-10 FROM N.A.
RX MEDLINE=90108722; PubMed=2691339;
RA Burt D.W., Mullins L.J., George H., Smith G., Brooks J., Pioli D., Brammar W.J.;
RT "The nucleotide sequence of a mouse renin-encoding gene, Ren-1d, and

RT its upstream region.";
RL Gene 84:91-104(1989).
RN [6]
RP SEQUENCE OF 1-30 FROM N.A.
RX MEDLINE=85085936; PubMed=6392850;
RA Field L.J., Philbrick W.M., Howles P.N., Dickinson D.P.,
RA McGowan R.A., Gross K.W.;
RT "Expression of tissue-specific Ren-1 and Ren-2 genes of mice:
comparative analysis of 5'-proximal flanking regions.";
RL Mol. Cell. Biol. 4:2321-2331(1984).
RN [7]
RP SEQUENCE OF 267-292 FROM N.A.
RX MEDLINE=84057744; PubMed=6357783;
RA Panthier J.J., Rougeon F.;
RT "Kidney and submaxillary gland renins are encoded by two non-allelic
genes in Swiss mice.";
RL EMBO J. 2:675-678(1983).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=92301530; PubMed=1608447;
RA Dhanraj V., Dealwis C.G., Frazao C., Badasso M., Sibanda B.L.,
RA Tickle I.J., Cooper J.B., Driessen H.P.C., Newman M., Aguilar C.,
RA Wood S.P., Blundell T.L., Hobart P.M., Geoghagan K.F., Ammirati M.J.,
RA Danley D.E., O'Connor B.A., Hoover D.J.;
RT "X-ray analyses of peptide-inhibitor complexes define the structural
basis of specificity for human and mouse renins.";
RL Nature 357:466-472(1992).
CC -!- FUNCTION: Renin is a highly specific endopeptidase, related to
pepsin, whose only known function is to generate angiotensin I
from angiotensinogen in the plasma, initiating a cascade of
reactions that produce an elevation of blood pressure and
increased sodium retention by the kidney. Its function in the
salivary gland is not understood.
CC -!- CATALYTIC ACTIVITY: Cleaves leu-|- bond in angiotensinogen to
generate angiotensin I.
CC -!- SUBUNIT: Dimer of a heavy chain and a light chain joined by a
disulfide bond.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Submandibular gland.
CC -!- MISCELLANEOUS: The active enzyme isolated from the submandibular
gland has catalytic and antigenic activities similar to renal
renin.
CC -!- SIMILARITY: Belongs to peptidase family A1.
CC -!- CAUTION: Ref.2 sequence differs from that shown in having 195-Leu-
Ser-Arg-Ser-198, which is due to a shift in the translation
reading frame, and Val-395. The authors' translation for 99 agrees
with that shown but does not agree with the nucleic acid sequence.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; J00621; AAA40050.1; ALT FRAME.
DR EMBL; BC011157; RAH11157.1; -.
DR EMBL; K02597; AAA40048.1; -.
DR EMBL; M34191; AAA40046.1; -.
DR EMBL; AF237860; AAA40047.1; -.
DR PIR; A93923; REMSS.
DR PIR; I77411; I77411.
DR PDB; 1SMR; 31-JAN-94.
DR MEROPS; A01.008; -.
DR MGD; MGI:97899; Ren2.
DR InterPro; IPR001969; Asparticase AS.
DR InterPro; IPR009007; Pept_A acid.
DR InterPro; IPR001461; Peptidase_A1.
DR Pfam; PF00026; asp; 1.
DR PRINTS; PR00792; PEPSIN.
DR PROSITE; PS00141; ASP PROTEASE; 2.
KW Hydrolase; Aspartyl protease; Plasma; Signal; Zymogen;

KW Submandibular gland; 3D-structure.
FT SIGNAL 1 25 PROBABLE.
FT PROPEP 26 63 ACTIVATION PEPTIDE.
FT CHAIN 64 401 RENIN 2.
FT CHAIN 64 351 RENIN 2.
FT CHAIN 354 401 RENIN 2 HEAVY CHAIN.
FT ACT_SITE 101 101 RENIN 2 LIGHT CHAIN.
FT ACT_SITE 286 286
FT DISULFID 114 121
FT DISULFID 277 281
FT DISULFID 320 357
FT CONFLICT 13 13 L -> W (IN REF. 5).
FT STRAND 69 69
FT STRAND 71 78
FT TURN 79 81
FT TURN 82 89
FT TURN 90 93
FT STRAND 94 101
FT TURN 102 103
FT STRAND 107 111
FT TURN 112 113
FT TURN 116 117
FT HELIX 119 123
FT STRAND 127 127
FT HELIX 129 131
FT TURN 133 134
FT STRAND 136 146
FT TURN 147 148
FT STRAND 149 162
FT TURN 163 164
FT STRAND 165 176
FT HELIX 179 182
FT TURN 183 184
FT STRAND 189 192
FT HELIX 196 198
FT TURN 200 202
FT HELIX 206 212
FT TURN 213 214
FT STRAND 216 216
FT STRAND 220 225
FT STRAND 234 238
FT HELIX 243 245
FT STRAND 246 254
FT STRAND 257 257
FT TURN 258 261
FT STRAND 262 270
FT TURN 271 272
FT STRAND 276 276
FT TURN 278 279
FT STRAND 281 285
FT TURN 287 288
FT STRAND 289 289
FT STRAND 292 294
FT HELIX 296 306
FT TURN 307 307
FT STRAND 309 312
FT TURN 313 314
FT STRAND 315 319
FT HELIX 320 325
FT STRAND 329 333
FT TURN 334 335
FT STRAND 336 340
FT HELIX 342 345
FT STRAND 346 346

Query Match 1.2%, Score 7; DB 1; Length 401;
Best Local Similarity 100.0%; Pred.No.74; 0; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;
QY 72 DXLCIVA 78
|||
Db 354 DXLCIVA 360

FT DISULFID 132 147 PROBABLE.
 FT VARIANT 14 14 H -> C (requires 2 nucleotide substitutions).
 FT VARIANT 101 101 A -> V.
 FT VARIANT 103 103 A -> T.
 FT VARIANT 257 257 D -> Y.
 FT VARIANT 377 377 A -> E.
 SQ SEQUENCE 405 AA; 45176 MW; CDDE31B82FEDEBE6 CRC64;
 Query Match 1.2%; Score 7; DB 1; Length 405;
 Best Local Similarity 100.0%; Pred.No. 75;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 568 FAEGKK 574
 DB 392 FAEGKK 398
 RESULT 58
 ID AUP1_MOUSE STANDARD; PRT; 410 AA.
 AC P70235;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ancient ubiquitous protein 1 precursor.
 GN AUP1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=96411699; PubMed=8812468;
 RA Jang W., Weber J.S., Bashir R., Bushby K., Meisler M.H.;
 RT "Aup1, a novel gene on mouse chromosome 6 and human chromosome 2p13.";
 RL Genomics 36:366-368(1996).
 CC -!- TISSUE SPECIFICITY: Ubiquitous.
 CC -!- SIMILARITY: Belongs to the AUP1 family.
 CC -!- FUNCTION: Belongs to the AUP1 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U41736; AAC52839.1; -.
 DR MGD; MGI:107789; Aup1.
 DR InterPro; IPR002123; Acyltransferase.
 DR Pfam; PF02845; CUE; 1.
 DR SMART; SM00546; CUE; 1.
 DR SMART; SM00563; PlsC; 1.
 KW Signal.
 FT SIGNAL 1 37 POTENTIAL.
 FT CHAIN 38 410 ANCIEN USIQUITOUS PROTEIN 1.
 SQ SEQUENCE 410 AA; 46121 MW; E7D070CBE296BD5B CRC64;
 Query Match 1.2%; Score 7; DB 1; Length 410;
 Best Local Similarity 100.0%; Pred.No. 76;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 331 LYEVARR 337
 DB 393 LYEVARR 399
 RESULT 59
 ID HS47_MOUSE STANDARD; PRT; 417 AA.

RESULT 57
 ID ARRS CANFA STANDARD; PRT; 405 AA.
 AC Q28281;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE S-arrestin (Retinal S-antigen) (48 kDa protein) (S-AG) (Rod
 DE photoreceptor arrestin).
 GN SAG OR SAG1 OR ARR.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Beagle X Briard; TISSUE=Retina;
 RX MEDLINE=97243904; PubMed=9088745;
 RA Veske A., Nafstroem K., Finckh U., Sargan D.R., Nilsson S.E.G.,
 RA Gal A.A.;
 RT "Isolation of canine retinal arrestin cDNA and exclusion of three
 RT candidate genes for Swedish Briard retinal dystrophy.";
 RL Curr. Eye Res. 16:270-274(1997).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANTS CYS-14; VAL-101; THR-103; TYR-257 AND
 GLU-377.
 RX MEDLINE=22325411; PubMed=12123530;
 RA Dekonien G., Eppien J.T.;
 RT "Screening of the arrestin gene in dogs afflicted with generalized
 RT progressive retinal atrophy.";
 RL BMC Genet. 3:12-12(2002).
 CC -!- FUNCTION: Arrestin is one of the major proteins of the rod
 CC (retinal rod outer segments); it binds to photoactivated-
 CC phosphorylated rhodopsin, thereby apparently preventing the
 CC transducin-mediated activation of phosphodiesterase.
 CC -!- TISSUE SPECIFICITY: Retina and pineal gland.
 CC -!- DISEASE: S-antigen induces autoimmune uveitis.
 CC -!- MISCELLANEOUS: Arrestin binds calcium (by similarity).
 CC -!- DISPAR: Defects in SAG may be the cause of generalized
 CC progressive retinal atrophy (GPRA) in some breeds.
 CC -!- SIMILARITY: Belongs to the arrestin family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X98460; CAA67100.1; -.
 DR EMBL; AJ426068; CAD19827.1; JOINED.
 DR EMBL; AJ426069; CAD19827.1; JOINED.
 DR EMBL; AJ426070; CAD19827.1; JOINED.
 DR EMBL; AJ426071; CAD19827.1; JOINED.
 DR EMBL; AJ426072; CAD19827.1; JOINED.
 DR EMBL; AJ426073; CAD19827.1; JOINED.
 DR EMBL; AJ426074; CAD19827.1; JOINED.
 DR EMBL; AJ426075; CAD19827.1; JOINED.
 DR EMBL; AJ426076; CAD19827.1; JOINED.
 DR EMBL; AJ426077; CAD19827.1; JOINED.
 DR EMBL; AJ426078; CAD19827.1; JOINED.
 DR HSSP; P08168; 1CF1.
 DR InterPro; IPR000658; Arrestin.
 DR InterPro; IPR007110; IG-like.
 DR Pfam; PF00339; arrestin; 1.
 DR Pfam; PF02752; arrestin_C; 1.
 DR PRINTS; PR00309; ARRESTIN.
 DR ProDom; PD002099; Arrestin; 1.
 DR PROSITE; PS00295; ARRESTINS; 1.
 KW Sensory transduction; Vision; Autoimmune uveitis; Calcium-binding;
 Polymorphism.

P19324;
 AC 01-NOV-1990 (Rel. 16, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DE 17-OCT-2001 (Rel. 40, Last annotation update)
 DE 47 kDa heat shock protein precursor (Collagen-binding protein 1)
 DE (Serine protease inhibitor J6).
 GN SERPINH1 OR CSPI OR HSP47.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 18-43.
 RC STRAIN=BALE/C;
 RX MEDLINE=92283255; PubMed=4317794;
 RA Takechi H., Hirayoshi K., Nakai A., Kudo H., Saga S., Kita T.,
 RA Nagata K.;
 RT "Molecular cloning of a mouse 47-kDa heat-shock protein (HSP47), a
 RT collagen-binding stress protein, and its expression during the
 RT differentiation of F9 teratocarcinoma cells.";
 RL Eur. J. Biochem. 206:323-329(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Teratocarcinoma;
 RX MEDLINE=90368798; PubMed=2394749;
 RA Wang S.-Y., Gudas L.J.;
 RT "A retinoic acid-inducible mRNA from F9 teratocarcinoma cells encodes
 RT a novel protease inhibitor homologue.";
 RL J. Biol. Chem. 265:15818-15822(1990).
 RN [3]
 RP ERRATUM
 RX MEDLINE=913110706; PubMed=1956236;
 RA Wang S.-Y., Gudas L.J.;
 RL J. Biol. Chem. 266:14135-14135(1991).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93246243; PubMed=8482533;
 RA Hosokawa N., Takechi H., Yokota S.I., Hirayoshi K., Nagata K.;
 RT "Structure of the gene encoding the mouse 47-kDa heat-shock protein
 RT (HSP47).";
 RL Gene 126:187-193(1993).
 CC -!- FUNCTION: Binds specifically to collagen. Could be involved as a
 CC chaperone in the biosynthetic pathway of collagen.
 CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
 CC -!- INDUCTION: By heat shock and retinoic acid.
 CC -!- SIMILARITY: Belongs to the serpin family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X60676; CAA43091.1; .
 CC EMBL; J05609; AAA03200.1; ALT_SEQ.
 CC EMBL; D12907; BAA02298.1; .
 CC EMBL; D12905; BAA02298.1; JOINED.
 CC EMBL; D12906; BAA02298.1; JOINED.
 CC F.R.; S23453; A42843.
 CC HSP; P05619; IHLB.
 CC MGD; MGI:88283; Serpinh1.
 CC GO; GO:0003773; F:heat shock protein activity; IDA.
 CC InterPro; IPR000886; ER target S.
 CC InterPro; IPR000215; Serpin.
 CC Pfam; PF00079; serpin; 1.
 CC SMART; SM00093; SERPIN; 1.
 CC PROSITE; PS00014; ER_TARGET; 1.
 CC PROSITE; PS00284; SERPIN; 1.
 KW Serpin; Signal; Heat shock; Endoplasmic reticulum; Glycoprotein;
 KW Chapterone.
 FT SIGNAL 1 17

FT CHAIN 18 417 47 kDa HEAT SHOCK PROTEIN.
 FT CARBOHYD 119 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 124 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 394 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT ACT_SITE 376 REACTIVE BOND (BY SIMILARITY).
 FT SITE 414 PREVENT SECRETION FROM ER (PROBABLE).
 FT CONFLICT 176 A -> P (IN REF. 2).
 FT CONFLICT 212 R -> K (IN REF. 4).
 FT CONFLICT 216 R -> K (IN REF. 4).
 FT CONFLICT 270 MP -> IA (IN REF. 2).
 FT CONFLICT 277 L -> S (IN REF. 2).
 SQ SEQUENCE 417 AA; 46589 MW; CD7C785EDC3C4F6A CRC64;
 Query Match 1.2%; Score 7; DB 1; Length 417;
 Best Local Similarity 100.0%; Pred. No. 77;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 540 TKEQLKA 546
 DB 285 TKEQLKA 291
 RESULT 60
 PROA_CLOAB STANDARD; PRT; 418 AA.
 AC 097E62;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Gamma-glutamyl phosphate reductase (GPR) (EC 1.2.1.41) (Glutamate-5-
 DE semialdehyde dehydrogenase) (Glutamyl-gamma-semialdehyde
 DE dehydrogenase) (GSA dehydrogenase).
 GN PROA OR CAC3254
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=11466286;
 RA Noelling J., Brston G., Onelchenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Wolf Y.I., Dally M.J.,
 RA Tsvetov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RT "Genome sequence and comparative analysis of the solvent-producing
 RT bacterium Clostridium acetobutylicum.";
 RL J. Bacteriol. 183:4823-4838(2001).
 CC -!- FUNCTION: Catalyzes the NADPH dependent reduction of L-gamma-
 CC glutamyl 5-phosphate into L-glutamate 5-semialdehyde and
 CC phosphate. The product spontaneously undergoes cyclization to form
 CC 1-pyrroline-5-carboxylate.
 CC -!- CATALYTIC ACTIVITY: L-glutamate 5-semialdehyde + phosphate +
 CC NADP(+) = L-gamma-glutamyl 5-phosphate + NADPH.
 CC -!- PATHWAY: Proline biosynthesis; second step.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the gamma-glutamyl phosphate reductase
 CC family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AE007821; AAK81188.1; .
 CC PIR; A97300; A97300.
 CC HAAAP; MF_00412; .
 CC InterPro; IPR002086; Aldehyde dehydr.
 CC InterPro; IPR000965; Gglut_pp_reduct.
 CC Pfam; PF00171; aldehyd; 1.

```
DR TIGRFAMs; TIGR00407; proA; 1.
DR PROSITE; PS01223; proA; 1.
KW Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
SQ SEQUENCE 418 AA; 45908 MW; 1F4AB047F388F46C CRC64;

Query Match 1.2%; Score 7; DB 1; Length 418;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 271 ISSKLKE 277
Db 278 ISSKLKE 284

RESULT 61
P47K_PSECL STANDARD; PRT; 419 AA.
AC P31521;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 47 kDa protein (P47K).
OS Pseudomonas chlororaphis (Pseudomonas aureofaciens).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=333;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B23;
RX MEDLINE=9119202; PubMed=2013568;
RA Nishiyama M., Horinouchi S., Kobayashi M., Nagasawa T., Yamada H.,
RA Beppu T.;
RT "Cloning and characterization of genes responsible for metabolism of
RT nitrite compounds from Pseudomonas chlororaphis B23.";
RL J. Bacteriol. 173:2465-2472(1991).
CC -!- FUNCTION: THE PRESENCE OF P47K IS CRITICAL FOR THE EXPRESSION
CC OF THE NITRILE HYDROLASE GENES. MAY STABILIZE OR ACTIVATE THE
CC NITRILE HYDROLASE PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D90216; BAA14247.1; -
DR PIR; D42725; D42725.
DR InterPro; IPR003495; Cobw.
DR Pfam; PF02492; cobw; 1.
SQ SEQUENCE 419 AA; 46666 MW; FF5113800E27FF0C CRC64;

Query Match 1.2%; Score 7; DB 1; Length 419;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 DLPSLAA 307
Db 233 DLPSLAA 239

RESULT 62
AUP1_HUMAN STANDARD; PRT; 476 AA.
AC Q9Y679; Q9UN06; Q9Y685;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ancient ubiquitous protein 1 precursor.
GN AUP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 297 EMPADLP 303
DB 111 EMPADLP 117

RESULT 64
POLG_HCVJ2
ID POLG_HCVJ2 STANDARD; PRT; 513 AA.
AC DT27959;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1)] (Fragment)
DS Hepatitis C virus (isolate HC-J2) (HCV)
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11111;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92230232; PubMed=1314459;
RA Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H.,
Tanaka T., Fukuda S., Tsuda F., Mishiro S.;
RT "Full-length sequence of a hepatitis C virus genome having poor
homology to reported isolates: comparative study of four distinct
genotypes."; J. Virol. 74:188-192 (1992).
RL Virology 188:331-341 (1992).
CC -!- FUNCTION: The small proteins NS2A, NS2B, NS4A and NS4B are
hydrophobic, suggesting a possible membrane-related function. NS3
and NS5 may play a role in the viral RNA replication.
CC -!- SUBUNIT: The virion of this virus is a nucleocapsid covered by a
lipoprotein envelope. The envelope consists of two proteins:
protein M and glycoprotein E. The nucleocapsid is a complex of
protein C and mRNA.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

EMBL; D10074; BAA00968.1; -
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR Pfam; PF01543; HCV_core; 1.
DR Pfam; PF01542; HCV_capsid; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR ProDom; PD186062; HCV_NS1; 1.
KW Polyprotein; Glycoprotein; Coat protein; Envelope protein;
KW Transmembrane; Nonstructural protein.
FT INIT_MET 1
FT CHAIN 1 115
FT CHAIN 116 191
FT CHAIN 192 383
FT CHAIN 384 >513
FT TRANSMEM 347 369
FT TRANSMEM 347 369
FT CARBOHYD 196 196
FT CARBOHYD 209 209
FT CARBOHYD 233 233
FT CARBOHYD 234 234
FT CARBOHYD 250 250
FT CARBOHYD 305 305
FT CARBOHYD 417 417
FT CARBOHYD 423 423

REMOVED FROM CAPSID PROTEIN C BY THE
CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
POTENTIAL.
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).

Pfam; PF02845; CUE; 1.
DR SMART; SMO0546; CUE; 1.
DR SMART; SMO0563; P18C; 1.
Signal; Alternative splicing.
SIGNAL 1 37
CHAIN 38 476
DOMAIN 160 165
VARSPLIC 114 179
CONFLICT 354 354
SEQUENCE 476 AA; 53028 MW; A6AF3A8D4AE86CB CRC64;
S -> P (IN REF. 3).
Query Match 1.2%; Score 7; DB 1; Length 476;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 331 LYEYARR 337
DB 459 LYEYARR 465

RESULT 63
TRPE_RHOSH
ID TRPE_RHOSH STANDARD; PRT; 500 AA.
AC P95646;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Anthranilate synthase component I (EC 4.1.1.3.27).
DS Rhodospirillum rubrum
OS Rhodospirillum rubrum
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Rhodobacter.
OX NCBI_TaxID=1063;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;
RA Rosanas A., Barbe J., Gilbert I.;
RT Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Chorismate + L-glutamine = anthranilate +
pyruvate + L-glutamate.
CC -!- PATHWAY: Tryptophan biosynthesis; first step.
CC -!- SUBUNIT: Tetramer of two components I and two components II (By
similarity).
CC -!- MISCELLANEOUS: Component I catalyzes the formation of anthranilate
using ammonia rather than glutamine, whereas component II provides
glutamine amidotransferase activity.
CC -!- SIMILARITY: Belongs to the anthranilate synthase component I
family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

EMBL; Y09072; CAA70293.1; -
DR HSP; Q06128; 10DL.
DR InterPro; IPR005801; Anth_synth_chor.
DR InterPro; IPR006805; Anth_synth_I_N.
DR InterPro; IPR005256; Anth_synth_I.
DR Pfam; PF04715; Anth_synth_I_N; 1.
DR Pfam; PF00425; chorismate_bind; 1.
DR PRINTS; PR00095; ANTSNTHASEI.
DR ProDom; PD000779; Anth_synth_chor; 1.
DR TIGRFAMs; TIGR00564; tpe_mocB; 1.
KW Tryptophan biosynthesis; Lyase.
SEQUENCE 500 AA; 54968 MW; 922945DF91F88F9 CRC64;
Query Match 1.2%; Score 7; DB 1; Length 500;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0

FT CARBOHYD 430 430 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 513 513
SQ SEQUENCE 513 AA; 55704 MW; 943F31E3514CDEF3 CRC64;

Query Match 1.2%; Score 7; DB 1; Length 513;
Best Local Similarity 100.0%; Pred.No. 92;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGKV 433
| | | | |
Db 116 SRNLGKV 122

RESULT 65
POLG HCVH4
ID POLG HCVH4 STANDARD; PRT; 520 AA.
AC Q01404;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polypeptide [Contains: Capsid protein C (Core protein) (P22);
DE Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
DE (GP68) (GP70) (NS1)] (Fragment).
OS Hepatitis C virus (isolate HCV-476) (HCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31643;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93019030; PubMed=1383400;
RA Abe K., Inchauspe G., Fujisawa K.;
RT "Genomic characterization and mutation rate of hepatitis C virus
RT isolated from a patient who contracted hepatitis during an epidemic
RT of non-A, non-B hepatitis in Japan."
RL J. Gen. Virol. 73:2725-2729 (1992).
CC -!- SUBUNIT: The virion of this virus is a nucleocapsid covered by a
CC lipoprotein envelope. The envelope consists of two proteins:
CC protein M and glycoprotein E. The nucleocapsid is a complex of
CC protein C and mRNA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; D10688; BAA01530.1; -
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR ProDom; PD186062; HCV NS1; 1.
KW Polyprotein; Glycoprotein; Coat protein; Envelope protein;
KW Transmembrane; Nonstructural protein.
FT INIT_MET 1
FT CHAIN 1 115
FT CHAIN 116 191
FT CHAIN 192 383
FT CHAIN 384 >520
FT TRANSMEM 347 369
FT CARBOHYD 186 196
FT CARBOHYD 209 209
FT CARBOHYD 234 234
FT CARBOHYD 305 305
FT CARBOHYD 418 418

FT CARBOHYD 424 424 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 431 431 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 449 449 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 520 520
SQ SEQUENCE 520 AA; 56499 MW; AAL35246CF20D525 CRC64;

Query Match 1.2%; Score 7; DB 1; Length 520;
Best Local Similarity 100.0%; Pred.No. 93;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGKV 433
| | | | |
Db 116 SRNLGKV 122

RESULT 66
POLG HCVHK
ID POLG HCVHK STANDARD; PRT; 520 AA.
AC Q01403;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polypeptide [Contains: Capsid protein C (Core protein) (P22);
DE Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
DE (GP68) (GP70) (NS1)] (Fragment).
OS Hepatitis C virus (isolate HCV-KF) (HCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31644;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93019030; PubMed=1383400;
RA Abe K., Inchauspe G., Fujisawa K.;
RT "Genomic characterization and mutation rate of hepatitis C virus
RT isolated from a patient who contracted hepatitis during an epidemic
RT of non-A, non-B hepatitis in Japan."
RL J. Gen. Virol. 73:2725-2729 (1992).
CC -!- SUBUNIT: The virion of this virus is a nucleocapsid covered by a
CC lipoprotein envelope. The envelope consists of two proteins:
CC protein M and glycoprotein E. The nucleocapsid is a complex of
CC protein C and mRNA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; D10687; BAA01529.1; -
DR PIR; JQ1925; JQ1925.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR ProDom; PD186062; HCV NS1; 1.
KW Polyprotein; Glycoprotein; Coat protein; Envelope protein;
KW Transmembrane; Nonstructural protein.
FT INIT_MET 1
FT CHAIN 1 115
FT CHAIN 116 191
FT CHAIN 192 383
FT CHAIN 384 >520
FT TRANSMEM 347 369
FT CARBOHYD 186 196
FT CARBOHYD 209 209
FT CARBOHYD 233 233

FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 424 424 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 431 431 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 449 449 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT NON TER 520 520
 SQ SEQUENCE 520 AA; 56476 MW; 1D2BD0A6FF27349B CRC64;

Query Match 1.2%; Score 7; DB 1; Length 520;
 Best Local Similarity 100.0%; Pred. No. 93;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGKV 433
 DB 116 SRNLGKV 122

RESULT 67
 C9B1_GLYEC STANDARD; PRT; 523 AA.
 AC P93149;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cytochrome P450 9B1 (EC 1.14.1.-) ((2S)-flavanone 2-hydroxylase)
 DE (licodione synthase) (Flavone synthase II) (CYP 9B1)
 GN CYP9B1.
 OS Glycyrrhiza echinata (Licorice).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Eurosid 1; Fabales; Fabaceae; Papilionoideae; Galegeae; Glycyrrhiza.
 OC NCBI_TaxID=46348;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Akashi T., Aoki T., Kameya N., Nakamura I., Ayabe S.-I.;
 RT "Two new cytochrome P450 cDNAs from elicitor-induced Licorice
 (Glycyrrhiza echinata L.) cells.";
 RL (In) Plant Gene Register PGR97-167.
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=98372652; PubMed=9708921;
 RA Akashi T., Aoki T., Ayabe S.-I.;
 RT "Identification of a cytochrome P450 cDNA encoding (2S)-flavanone
 2-hydroxylase of licorice (Glycyrrhiza echinata L.; Fabaceae) which
 represents licodione synthase and flavone synthase II.";
 RL FEBS Lett. 431:287-290(1998).
 CC -1- FUNCTION: Catalyzes the formation of [14C]licodione and [14C]2-
 hydroxynaringenin from (2S)-[14C]liquiritigenin and (2S)-
 [14C]naringenin, respectively. Can also convert eriodictyol to
 luteolin.
 CC -1- INDUCTION: By fungal elicitor.
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB001380; BAA22423.1; -
 CC HSP; P14779; IJUZ.
 CC InterPro; IPR001128; Cytochrome_P450.
 CC Pfam; PF00067; P450; 1.
 CC PRINTS; PR00385; P450.
 CC PROSITE; PS00086; CYTOCHROME_P450; 1.
 CC Oxidoreductase; Monooxygenase; Transmembrane; Heme.
 CC TRANSMEM 66 86
 CC FT METAL 448 448 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 523 AA; 59495 MW; 4A776E6C432F527C CRC64;

Query Match 1.2%; Score 7; DB 1; Length 523;
 Best Local Similarity 100.0%; Pred. No. 94;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 528 ALVELYK 534
 DB 316 ALVELYK 322

RESULT 68
 HO_YEAST STANDARD; PRT; 586 AA.
 AC P09332; Q12183;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Homoethallic switching endonuclease (Ho endonuclease).
 GN HO OR YDL227C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OC NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87089786; PubMed=3025649;
 RA Russell D.W., Jensen R., Zoller M.J., Burke J., Errede B., Smith M.,
 RA Herskowitz I.,
 RT "Structure of the Saccharomyces cerevisiae HO gene and analysis of
 RT its upstream regulatory region.";
 RL Mol. Cell. Biol. 6:4281-4294(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=96120866; PubMed=8590483;
 RA Meiron H., Nahon E., Raveh D.;
 RT "Identification of the heterothallic mutation in HO-endonuclease of
 RT S. cerevisiae using HO/ho chimeric genes.";
 RL Curr. Genet. 28:367-373(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Rasmussen S.W.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP REVIEW.
 RX MEDLINE=91171884; PubMed=2005783;
 RA Herskowitz I., Jensen R.;
 RT "Putting the HO gene to work: practical uses for mating-type
 RT switching.";
 RL Meth. Enzymol. 194:132-146(1991).
 RN [5]
 RP DEGRADATION.
 RX MEDLINE=20420335; PubMed=10963670;
 RA Kaplan L., Ivantsov Y., Kornitzer D., Raveh D.;
 RT "Functions of the DNA damage response pathway target Ho endonuclease
 RT of yeast for degradation via the ubiquitin-26S proteasome system.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:10077-10082(2000).
 CC -1- FUNCTION: Initiation of mating type interconversion. This protein
 CC is a site-specific endonuclease that cleaves a site in the mat
 CC locus on chromosome III. The double-strand break is followed by a
 CC unidirectional gene conversion event that replaces the information
 CC at the mat locus by information copied from either of the two
 CC homologous loci (HMR and HML) that reside at the extremity of the
 CC chromosome III. Endonuclease expression takes place in late G1
 CC just before cells enter S phase.
 CC -1- FIM: Rapidly degraded via the ubiquitin-26S proteasome system
 CC through two ubiquitin-conjugating enzymes UBC2/RAD6 and
 CC UBC3/CDC34.
 CC -1- MISCELLANEOUS: The metal-binding domain form zinc-fingers that are
 CC involved in binding of the DNA.
 CC -1- SIMILARITY: TO YEAST VMA1-DERIVED ENDONUCLEASE (VDE).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M14678; AAA34683.1; -;
DR EMBL; X90957; CAA62447.1; -;
DR EMBL; Z74275; CAA98806.1; -;
DR PIR; S59301; S59301
DR HSSP; P17255; 1VDS
DR GenOnline; 140470; -;
DR SGB; S0002386; HO
DR GO; GO:0004519; F:endonuclease activity; IDA.
DR GO; GO:0007533; P:maturing-type switching/recombination; IMP.
DR InterPro; IPR003587; Hedgehog_hint_N.
DR InterPro; IPR007869; Hom_end.
DR InterPro; IPR007868; Hom_end_hint.
DR InterPro; IPR006142; INTEN_endonuc.
DR InterPro; IPR004042; Intein_endonuc.
DR Pfam; PF05204; Hom_end; 1.
DR Pfam; PF05203; Hom_end_hint; 1.
DR PRINTS; PR00379; INTEN.
DR SMART; SM00306; HintN; 1.
DR PROSITE; PS00819; INTEN_ENDONUCLEASE; 1.
KW Hydrolase; Endonuclease; Zinc; Zinc-finger; DNA-binding.
FT DOMAIN 215 370
FT CONFLICT 189 189 A -> T (IN REF. 1).
FT CONFLICT 223 223 S -> G (IN REF. 1).
FT CONFLICT 405 405 S -> L (IN REF. 1).
FT CONFLICT 475 475 L -> H (IN REF. 1).
SQ SEQUENCE 586 AA; 66089 MW; 95771394D177823A CRC64;

Query Match 1.2%; Score 7; DB 1; Length 586;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 KPLELEKS 287
DB 94 KPLELEKS 100
|||||

RESULT 69
DNAA MYCCA
ID DNAA MYCCA STANDARD; PRT; 591 AA.
AC P45958;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chaperone protein dnaK (Heat shock protein 70) (Heat shock 70 kDa protein) (HSP70).
GN DNAA.
OS Mycoplasma capricolum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2095;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25416;
RX MEDLINE=97148974; PubMed=8995799;
RA Falah M., Gupta R.S.;
RT "Phylogenetic analysis of mycoplasmas based on Hsp70 sequences: cloning of the dnaK (hsp70) gene region of Mycoplasma capricolum."; Int. J. Syst. Bacteriol. 47:38-45(1997).
RL [2]
RN SEQUENCE OF 1-227 FROM N.A.
RC STRAIN=ATCC 27343 / Kid;
RA Gillevet P., Ally A., Barton F., Brenner S.E., Clark-Whitehead R., Dolan M., Douglas N., Heu E., Purzcki M.S., Richter B., Russo S., Sartell J., Smith S.W., Wang C., Williams J., Gilbert W.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Acts as a chaperone (By similarity).
CC -!- INDUCTION: By stress conditions e.g. heat shock (By similarity).
CC -!- SIMILARITY: Belongs to the heat shock protein 70 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U51235; AAB09430.1; -;
DR EMBL; Z33106; CAA83764.1; -;
DR PIR; S7870; S7870.
DR HSSP; P04475; IDG4.
DR HAVAP; MF_00332; -; 1.
DR InterPro; IPR001023; Hsp70.
DR Pfam; PF00012; HSP70; 1.
DR PRINTS; PR00301; HEATSHOCK70.
DR ProDom; PD000089; Hsp70; 1.
DR PROSITE; PS00297; HSP70_1; 1.
DR PROSITE; PS00329; HSP70_2; 1.
DR PROSITE; PS01036; HSP70_3; 1.
KW Chaperone; ATP-binding; Heat shock; Phosphorylation.
FT MOD RES 175 175 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 591 AA; 63869 MW; B962340066F52343 CRC64;

Query Match 1.2%; Score 7; DB 1; Length 591;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 317 KNVAAEK 323
DB 100 KNVAAEK 106
|||||

RESULT 70
AFAM HUMAN
ID AFAM HUMAN STANDARD; PRT; 599 AA.
AC P43652;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Afamin precursor (Alpha-albumin) (Alpha-Alb).
GN AFM OR ALBA OR ALB2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=94299534; PubMed=7517938;
RA Lichtenstein H.S., Lyons D.E., Wurfel M.M., Johnson D.A., McGinley M.D., Leidl J.C., Trollinger D.B., Mayer J.P., Wright S.D., Zukowski M.M.;
RT "Afamin is a new member of the albumin, alpha-fetoprotein, and vitamin D-binding protein gene family."; J. Biol. Chem. 269:18149-18154(1994).
RL [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=96240683; PubMed=8648639;
RA Nishio H., Heiskanen M., Palotie A., Belanger L., Dugaiczky A.;
RT "Tandem arrangement of the human serum albumin multigene family in the sub-centromeric region of 4q: evolution and chromosomal direction of transcription."; J. Mol. Biol. 259:113-119(1996).
RL [3]
RN SEQUENCE OF 1-69; 105-207 AND 560-599 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=95180738; PubMed=7875606;
RA Allard D., Gilbert S., Lamontagne A., Hamel D., Belanger L.;
RT "Identification of rat alpha-albumin and cDNA cloning of its human ortholog."; Gene 153:287-288(1995).
RL

```

Query Match          1.2%; Score 7; DB 1; Length 599;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          268 QDSISSK 274
              |||||
Db          292 QDSISSK 298

```

KW Sodium transport; Transmembrane; Glycoprotein; Signal; Repeat;
KW Alternative splicing;
FT SIGNAL 21
FT CHAIN 22 605
FT DOMAIN 22 80
FT TRANSMEM 81 101
FT DOMAIN 102 155
FT TRANSMEM 156 176
FT DOMAIN 177 183
FT TRANSMEM 184 204
FT DOMAIN 205 207
FT TRANSMEM 228 228
FT DOMAIN 229 407
FT TRANSMEM 408 428
FT DOMAIN 429 440
FT TRANSMEM 441 461
FT DOMAIN 462 462
FT TRANSMEM 483 483
FT DOMAIN 484 509
FT TRANSMEM 510 530
FT DOMAIN 531 540
FT TRANSMEM 541 561
FT DOMAIN 562 569
FT TRANSMEM 570 590
FT DOMAIN 591 605
FT REPEAT 122 162
FT REPEAT 478 509
FT DOMAIN 378 389
FT CARBOHYD 52 52
FT CARBOHYD 59 59
FT VARSPLIC 1 47
FT VARSPLIC 1 92
FT VARSPLIC 155 193
FT VARSPLIC 258 276
FT VARSPLIC 458 483
FT FT
FT FT
FT VARSPLIC 484 605
FT FT
FT CONFLICT 427 427
FT SEQUENCE 605 AA; 56951 MW; E032612F61FA8E1A CRC64;
Query Match 1.2%; Score 7; DB 1; Length 605;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 341 YSVWLL 347
DB 543 YSVWLL 549
RESULT 72
ID NKX4 MOUSE STANDARD; PRT; 605 AA.
AC Q8CG08; Q8BLL5; Q8BLL7;
DT 10-OCT-2003 (Rel. 42; Created)
DT 10-OCT-2003 (Rel. 42; Last sequence update)
DE 10-OCT-2003 (Rel. 42; Last annotation update)
DE Sodium/potassium/calcium exchanger 4 precursor (Na(+)/K(+)/Ca(2+)-exchange protein 4).
GN SLC24A4 OR NKX4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=CS7BL/6J;
RC MEDLINE=22359087; PubMed=12379639;
RX

RA Li X.-P., Kraev A.S., Lytton J.;
RT "Molecular cloning of a fourth member of the potassium-dependent sodium-calcium exchanger gene family, NKX4.";
RL J. Biol. Chem. 277:48410-48417(2002).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2), AND SEQUENCE OF 1-240 FROM N.A.
RX (ISOFORM 1).
RC STRAIN=CS7BL/6J; TISSUE=Retina;
RC MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Nikaide I., Osato N., Saito R., Suzuki H., Yamahaka I., Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schraml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S., Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S., Gasterland T., Gariboldi M., Gissi C., Godzik A., Gough J., Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D., Kanai A., Kawai H., Kawasawa Y., Redzierski R.M., King B.B., Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A., Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H., Nagashima T., Numata K., Okido T., Pavan W.J., Perlea G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M., Sandelin A., Schneider C., Sempie C.A., Setou M., Shimada K., Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C., Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L., Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I., Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
CC -!- FUNCTION: Transports one Ca(2+) and one K(+) in exchange for four Na(+).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q8CG08-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q8CG08-2; Sequence=VSP_008375, VSP_008376, VSP_008377, VSP_008378;
CC -!- TISSUE SPECIFICITY: Widely expressed in most regions of the brain, including hippocampus, neocortex, thalamus, striatum and olfactory bulb.
CC -!- SIMILARITY: BELONGS TO THE SLC24A FAMILY OF TRANSPORTERS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL; AY156046; AA037415.1; -;
CC EMBL; AK044239; BAC31835.1; -;
CC EMBL; AK044368; BAC31887.1; ALT_INIT.
CC MGD; MGI:2447362; SLC24a4.
CC GO; GO:0005887; C: integral to plasma membrane; IC.
CC InterPro; IPR004481; K: NaCaexchng.
CC InterPro; IPR004937; NaCa_Ex; 2.
CC Pfam; PF01699; NaCa_Ex; 2.
CC TIGRFAMs; TIGR00367; TIGR00367; 1.
KW Transport; Antiport; Symport; Calcium transport; Potassium transport;
KW Sodium transport; Transmembrane; Glycoprotein; Signal; Repeat;
KW Alternative splicing.

FT SIGNAL 1 21
 FT CHAIN 22 605
 FT DOMAIN 22 80
 FT TRANSMEM 81 101
 FT TRANSMEM 102 155
 FT TRANSMEM 156 176
 FT TRANSMEM 177 183
 FT TRANSMEM 184 204
 FT TRANSMEM 205 207
 FT TRANSMEM 208 228
 FT TRANSMEM 229 407
 FT TRANSMEM 408 428
 FT TRANSMEM 429 440
 FT TRANSMEM 441 461
 FT TRANSMEM 462 462
 FT TRANSMEM 463 483
 FT TRANSMEM 484 509
 FT TRANSMEM 510 530
 FT TRANSMEM 531 540
 FT TRANSMEM 541 561
 FT TRANSMEM 562 569
 FT TRANSMEM 570 590
 FT TRANSMEM 591 605
 FT REPEAT 122 162
 FT REPEAT 478 509
 FT DOMAIN 378 387
 FT CARBOHYD 52 52
 FT CARBOHYD 59 59
 FT VARSPLIC 1 47
 FT VARSPLIC 258 276
 FT VARSPLIC 557 566
 FT VARSPLIC 567 605
 FT CONFLICT 413 413
 FT SEQUENCE 605 AA; 66866 MW; A32D256278FB0686 CRC64;

Query Match 1.2%; Score 7; DB 1; Length 605;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 YSVVLL 347
 DB 543 YSVVLL 549

RESULT 73
 AFAM RAT
 ID AFAM RAT STANDARD; PRT; 608 AA.
 AC P36953;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Afamin precursor (Alpha-albumin) (Alpha-Alb).
 GN AFM.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=94164881; PubMed=7503788;
 RA Belanger L., Roy S., Allard D.;
 RT "New albumin gene 3, adjacent to the alpha 1-fetoprotein locus.";
 RL J. Biol. Chem. 269:5481-5484(1994).
 CC -!- FUNCTION: Possible role in the transport of yet unknown ligand.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
 CC -!- SIMILARITY: Contains 3 albumin domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; X76456; CAAS3994.1; -
 CC PIR; A53195; A53195.
 CC HSSP; P02768; 1E7B.
 CC InterPro; IPR000264; Serum albumin.
 CC Pfam; PF00273; transport_prot; 3.
 CC PRINTS; PRO0802; SERUMALBUMIN.
 CC ProDom; PD002486; Serum_albumin; 1.
 CC SMART; SM00103; ALBUMIN; 3.
 CC PROSITE; PS00212; ALBUMIN; 1.
 CC Transport; Repeat; Glycoprotein; Signal.
 CC SIGNAL 1 21
 CC CHAIN 21 608
 CC DOMAIN 22 205
 CC DOMAIN 212 397
 CC DOMAIN 404 593
 CC DISULFID 77 86
 CC DISULFID 99 114
 CC DISULFID 113 124
 CC DISULFID 148 193
 CC DISULFID 224 270
 CC DISULFID 289 303
 CC DISULFID 302 313
 CC DISULFID 340 385
 CC DISULFID 384 393
 CC DISULFID 416 462
 CC DISULFID 451 470
 CC DISULFID 483 499
 CC DISULFID 498 509
 CC DISULFID 580 589
 CC CARBOHYD 33 33
 CC CARBOHYD 109 109
 CC CARBOHYD 153 153
 CC CARBOHYD 402 402
 CC CARBOHYD 488 488
 CC SEQUENCE 608 AA; 69335 MW; F33151A5E568A07F6 CRC64;

Query Match 1.2%; Score 7; DB 1; Length 608;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 268 QDSISK 274
 DB 292 QDSISK 298

RESULT 74
 AFAM MOUSE
 ID AFAM MOUSE STANDARD; PRT; 611 AA.
 AC O89020;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Afamin precursor (Alpha-albumin) (Alpha-Alb).
 GN AFM.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=diaphragm;
 RA van Reeth T., Gabant P., Dreze P., Szpirer J., Szpirer C.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Possible role in the transport of yet unknown ligand.

CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ011080; CAA09471.1; -.
DR HSP; P02768; 1E7B.
DR MGD; MGI:2429409; Afm.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_pro; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR PRODOM; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
KW Transport; Repeat; Glycoprotein; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 1 21 AFAMIN.
FT DOMAIN 22 611 ALBUMIN 1.
FT DOMAIN 212 397 ALBUMIN 2.
FT DOMAIN 404 593 ALBUMIN 3.
FT DISULFID 77 86 BY SIMILARITY.
FT DISULFID 99 114 BY SIMILARITY.
FT DISULFID 113 124 BY SIMILARITY.
FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 340 385 BY SIMILARITY.
FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 470 BY SIMILARITY.
FT DISULFID 483 499 BY SIMILARITY.
FT DISULFID 498 509 BY SIMILARITY.
FT DISULFID 580 589 BY SIMILARITY.
FT CARBOHYD 33 33 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 109 109 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 153 153 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 488 488 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 611 AA; 69635 MW; 39E46B6E723F89C8 CRC64;

Query Match 1.2%; Score 7; DB 1; Length 611;
Best Local Similarity 100.0%; Pred. No. 1.1e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 268 QDSISSK 274
| | | | |
Db 292 QDSISSK 298

RESULT 75
ID VP3B_HUMAN STANDARD; PRT; 617 AA.
AC Q9H267; Q96K14; Q9NRP6; Q9NSF3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Vacuolar protein sorting 33B (hVPS33B).
GN VPS33B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=20354999; PubMed=10894945;
RA Carim L., Sumoy L., Andreu N., Estivill X., Escarceller M.;
RT "Cloning, mapping and expression analysis of VPS33B, the human
RT orthologue of rat Vps33b.";
RL Cytogenet. Cell Genet. 89:92-95(2000).
RV [2]
RP SEQUENCE FROM N.A. AND VARIANT GLY-514.
RX MEDLINE=21147934; PubMed=11250079;
RA Huizing M., Didier A., Walenta J., Anikster Y., Gahl W.A., Kraemer H.;
RT "Molecular cloning and characterization of human VPS18, VPS16,
RT and VPS33.";
RL Gene 264:241-247(2001).
RV [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
RT "NED0 human cDNA sequencing project.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RV [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore I., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Faneay J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smaluk D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: May play a role in vesicle-mediated protein trafficking
CC to lysosomal compartments and in membrane docking/fusion reactions
CC of late endosomes/lysosomes (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic; peripheral membrane protein
CC associated with late endosomes/lysosomes (By similarity).
CC -!- TISSUE SPECIFICITY: Ubiquitous; highly expressed in testis and low
CC expression in the lung.
CC -!- SIMILARITY: Belongs to the STXBP/UNC-18/SEC1 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF201694; AAF91174.1; -.
DR EMBL; AL357472; CAB93109.1; -.
DR EMBL; AF308803; AAG34680.1; -.
DR EMBL; AK027754; BAB55345.1; -.
DR EMBL; BC016445; AAI16445.1; -.
DR Genbank; HGNC:12712; VPS33B.
DR InterPro; IPR001619; Sec1-like.
DR Pfam; PF00995; Sec1; 1.
KW Transport; Protein transport; Membrane; Polymorphism.
RV VARIANT 514 514 S -> G.

FT CONFLICT 293 293 /FTid=VAR_013828.
FT CONFLICT 466 466 H -> Y (IN REF. 3).
SQ SEQUENCE 617 AA; 7061S MW; BC69B83B0BAA192E CRC64;
K -> E (IN REF. 3).
Query Match 1.2%; Score 7; DB 1; Length 617;
Best Local Similarity 100.0%; Pred.No. 1.1e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 231 VSKLVTD 237
Db 459 VSKLVTD 465

Search completed: April 19, 2004, 16:06:18
Job time : 20 secs

GenCore version 5.1.6		Copyright (c) 1993 - 2004 Compugen Ltd.		OM protein - protein search, using sw model		Run on: April 19, 2004, 16:04:40 ; Search time 22 Seconds (without alignments)		1372.782 Million cell updates/sec		Title: US-09-832-929-18		Perfect score: 585		Sequence: 1 DAHKSEVHRFKDLGEBNFK.....TCFABEGKKLVAAASQAALGL 585		Scoring table: OLIGO		Gapop 60.0 , Gapext 60.0		Searched: 389414 seqs, 51625971 residues		Word size : 0		Total number of hits satisfying chosen parameters: 389414		Minimum DB seq length: 0		Maximum DB seq length: 2000000000		Post-processing: Listing first 100 summaries		Database :		Issued Patents AA:*		1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*		2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*		3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*		4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*		5: /cgn2_6/ptodata/2/iaa/PTUS COMB.pep.*		6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*		Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.		SUMMARIES		Result No. Score Query Match Length DB ID Description		1 585 100.0 585 1 US-08-153-799-14 Sequence 14, Appl		2 585 100.0 585 2 US-08-702-572-2 Sequence 2, Appl		3 585 100.0 585 3 US-08-759-746-2 Sequence 5, Appl		4 585 100.0 585 4 US-10-153-064-5 Sequence 7, Appl		5 585 100.0 609 4 US-09-976-594-977 Sequence 977, App		6 585 100.0 610 2 US-08-797-689-2 Sequence 2, Appl		7 585 100.0 610 4 US-09-984-186-2 Sequence 133, App		8 585 100.0 651 4 US-10-153-064-133 Sequence 132, App		9 585 100.0 652 4 US-10-153-064-132 Sequence 131, App		10 585 100.0 653 4 US-10-153-064-131 Sequence 130, App		11 585 100.0 656 4 US-10-153-064-130 Sequence 129, App		12 585 100.0 676 4 US-10-153-064-127 Sequence 127, App		13 585 100.0 676 4 US-10-153-064-129 Sequence 125, App		14 585 100.0 677 4 US-10-153-064-125 Sequence 123, App		15 585 100.0 680 4 US-10-153-064-123 Sequence 2, Appl		16 585 100.0 783 1 US-08-256-938-2 Sequence 4, Appl		17 585 100.0 787 1 US-08-256-938-4 Sequence 16, Appl		18 585 100.0 787 2 US-08-797-689-16 Sequence 16, Appl		19 585 100.0 787 4 US-09-984-186-16 Sequence 96, Appl		20 585 100.0 787 4 US-10-153-064-96 Sequence 99, Appl		21 510 87.2 652 4 US-10-153-064-96 Sequence 105, App		22 510 87.2 652 4 US-10-153-064-99 Sequence 90, Appl		23 510 87.2 652 4 US-10-153-064-105 Sequence 93, Appl		24 510 87.2 660 4 US-10-153-064-90 Sequence 95, Appl		25 510 87.2 660 4 US-10-153-064-93 Sequence 98, Appl		26 510 87.2 676 4 US-10-153-064-95 Sequence 98, Appl		27 510 87.2 676 4 US-10-153-064-98 Sequence 98, Appl		28 510 87.2 676 4 US-10-153-064-104 Sequence 92, Appl		29 510 87.2 684 4 US-10-153-064-92 Sequence 89, Appl		30 484 82.7 609 1 US-08-222-619-3 Sequence 3, Appl		31 484 82.7 609 1 US-08-433-037-4 Sequence 4, Appl		32 484 82.7 609 1 US-08-897-956A-2 Sequence 2, Appl		33 484 82.7 609 5 PCT-US95-04075-3 Sequence 3, Appl		34 484 82.7 609 4 US-08-897-956A-3 Sequence 102, App		35 483 82.6 668 4 US-10-153-064-102 Sequence 101, App		36 478 81.7 692 4 US-10-153-064-101 Sequence 3, Appl		37 383 65.5 585 1 US-08-448-196A-3 Sequence 1, Appl		38 383 65.5 585 2 US-09-846-329A-1 Sequence 1, Appl		39 27 4.6 28 4 US-08-846-329A-1 Sequence 5, Appl		40 27 4.6 583 1 US-08-448-196A-5 Sequence 7, Appl		41 22 3.8 582 1 US-08-448-196A-7 Sequence 1, Appl		42 22 3.8 582 1 US-08-134-638-1 Sequence 4, Appl		43 22 3.8 583 1 US-08-448-196A-4 Sequence 6, Appl		44 22 3.8 583 1 US-08-448-196A-6 Sequence 2, Appl		45 22 3.2 117 1 US-08-470-187-1 Sequence 1, Appl		46 17 2.9 17 1 US-08-318-905-1 Sequence 1, Appl		47 17 2.9 17 1 US-08-483-232-1 Sequence 1, Appl		48 17 2.9 17 1 US-08-483-140-1 Sequence 1, Appl		49 17 2.9 17 2 US-08-485-938A-1 Sequence 1, Appl		50 17 2.9 17 2 US-08-910-041-1 Sequence 1, Appl		51 17 2.9 17 3 US-09-328-474-1 Sequence 1, Appl		52 17 2.9 17 3 US-09-100-546-1 Sequence 1, Appl		53 17 2.9 17 3 US-09-010-715-1 Sequence 1, Appl		54 17 2.9 17 3 US-09-577-758-1 Sequence 16, Appl		55 17 2.6 15 2 US-08-702-572-16 Sequence 24, Appl		56 15 2.6 16 4 US-10-053-485-24 Sequence 30, Appl		57 13 2.2 13 1 US-08-787-547-30 Sequence 7, Appl		58 13 2.2 13 2 US-08-803-364-7 Sequence 13, Appl		59 13 2.2 13 2 US-09-024-198-13 Sequence 13, Appl		60 13 2.2 13 2 US-09-186-409-13 Sequence 10, Appl		61 12 2.1 12 1 US-08-153-799-10 Sequence 26, Appl		62 12 2.1 17 4 US-10-053-485-26 Sequence 4, Appl		63 11 1.9 11 1 US-08-378-859-4 Sequence 4, Appl		64 11 1.9 11 3 US-08-970-648-4 Sequence 1, Appl		65 11 1.9 11 4 US-08-952-558-1 Sequence 25, Appl		66 11 1.9 16 4 US-10-053-485-25 Sequence 1, Appl		67 10 1.7 10 3 US-08-378-859-1 Sequence 1, Appl		68 10 1.7 10 3 US-08-970-648-1 Sequence 2, Appl		69 10 1.7 10 4 US-08-952-558-2 Sequence 9, Appl		70 9 1.5 11 1 US-08-469-856-9 Sequence 3, Appl		71 8 1.4 8 3 US-08-378-859-3 Sequence 3, Appl		72 8 1.4 8 4 US-09-165-926-2 Sequence 2, Appl		73 8 1.4 8 4 US-09-165-961-2 Sequence 2, Appl		74 8 1.4 8 4 US-08-165-581-2 Sequence 4, Appl		75 8 1.4 8 4 US-08-352-558-4 Sequence 21, Appl		76 8 1.4 12 3 US-10-053-485-21 Sequence 13, Appl		77 8 1.4 12 3 US-08-477-928A-11 Sequence 11, Appl		78 8 1.4 18 3 US-08-477-928A-12 Sequence 12, Appl		79 8 1.4 197 3 US-08-377-309-3 Sequence 3, Appl		80 8 1.4 197 3 US-08-186-723-3 Sequence 6, Appl		81 8 1.4 197 4 US-09-186-949A-4 Sequence 4, Appl		82 8 1.4 197 4 US-08-758-757-3 Sequence 3, Appl		83 8 1.4 197 4 US-09-187-978-3 Sequence 3, Appl		84 8 1.4 197 4 US-10-115-701A-3 Sequence 3, Appl		85 8 1.4 197 5 PCT-US96-0098-6 Sequence 6, Appl		86 8 1.4 362 2 US-08-377-309-6 Sequence 6, Appl		87 8 1.4 389 2 US-08-186-723-6 Sequence 9, Appl		88 8 1.4 389 4 US-08-505-012-9 Sequence 7, Appl		89 8 1.4 389 4 US-09-186-949A-7 Sequence 6, Appl		90 8 1.4 389 4 US-08-187-978-6 Sequence 6, Appl		91 8 1.4 389 4 US-10-115-701A-6 Sequence 6, Appl		92 8 1.4 389 4 US-10-115-701A-6 Sequence 6, Appl		93 8 1.4 389 4 US-10-115-701A-6 Sequence 6, Appl		94 8 1.4 389 4 US-10-115-701A-6 Sequence 6, Appl		95 8 1.4 389 4 US-10-115-701A-6 Sequence 6, Appl		96 8 1.4 389 4 US-10-115-701A-6 Sequence 6, Appl		97 8 1.4 389 4 US-10-115-701A-6 Sequence 6, Appl		98 8 1.4 389 4 US-10-115-701A-6 Sequence 6, Appl		99 8 1.4 389 4 US-10-115-701A-6 Sequence 6, Appl		100 8 1.4 389 4 US-10-115-701A-6 Sequence 6, Appl	
-----------------------	--	---	--	---	--	--	--	-----------------------------------	--	-------------------------	--	--------------------	--	---	--	----------------------	--	--------------------------	--	--	--	---------------	--	---	--	--------------------------	--	-----------------------------------	--	--	--	------------	--	---------------------	--	--	--	--	--	--	--	--	--	--	--	---	--	--	--	-----------	--	---	--	--	--	--	--	--	--	--	--	---	--	--	--	---	--	---	--	---	--	--	--	--	--	--	--	--	--	--	--	---	--	---	--	--	--	---	--	---	--	---	--	--	--	--	--	---	--	--	--	--	--	--	--	--	--	---	--	--	--	--	--	--	--	---	--	---	--	--	--	---	--	--	--	---	--	---	--	--	--	---	--	---	--	--	--	---	--	---	--	--	--	---	--	---	--	---	--	--	--	---	--	---	--	---	--	---	--	--	--	---	--	---	--	--	--	--	--	---	--	---	--	---	--	--	--	---	--	---	--	--	--	--	--	---	--	---	--	---	--	--	--	---	--	---	--	---	--	---	--	--	--	--	--	---	--	---	--	---	--	---	--	--	--	---	--	---	--	--	--	---	--	---	--	---	--	---	--	--	--	---	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	---	--

ALIGNMENTS

RESULT 1
US-08-153-799-14
; Sequence 14, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Region
; LOCATION: 369..419
; OTHER INFORMATION: /note= "Alternative C-termini of
; OTHER INFORMATION: HSA(1-n)"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..585
; OTHER INFORMATION: /note= "Amino acid sequence of
; OTHER INFORMATION: natural HSA"
; US-08-153-799-14
; Query Match 100.0%; Score 585; DB 1; Length 585;
; Best Local Similarity 100.0%; Pred. No. 0;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRFKDGLSENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60
DB 1 DAHKSEVAHRFKDGLSENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPPLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPPLRLVRPEV 120
QY 121 DVNCTAFHDNEETFLKXLYIETARRHPYFAPBELLFFAKRYKAAFTCCOAAADKAACLLP 180
DB 121 DVNCTAFHDNEETFLKXLYIETARRHPYFAPBELLFFAKRYKAAFTCCOAAADKAACLLP 180
QY 181 KLDELDEGKASSAKORLKASLQKGFGERAFKAWAVARLSQRFPAEFAEVSCLVTDLT 240
DB 181 KLDELDEGKASSAKORLKASLQKGFGERAFKAWAVARLSQRFPAEFAEVSCLVTDLT 240
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISKLEKCEKPLLEKSHCIAEVENDENPA 300
DB 241 VHTCCCHGDLLECADRADLAKYICENQDSISKLEKCEKPLLEKSHCIAEVENDENPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEARHDPDYVLLLLRLAKTYETTTLEKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEARHDPDYVLLLLRLAKTYETTTLEKC 360
QY 361 CAAADPHECYAKVDFEFKPLVEBPQNLKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
DB 361 CAAADPHECYAKVDFEFKPLVEBPQNLKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETVVPKEFNAETFTFHADICTLSEKERQIKQOTALVELVGHKPKAT 540
DB 481 LVNRRPCFSALEVDETVVPKEFNAETFTFHADICTLSEKERQIKQOTALVELVGHKPKAT 540
QY 541 KEQLKAVMDDFAAFVKECKCKADDKTCFAEEGKGLVAASQAALGL 585
DB 541 KEQLKAVMDDFAAFVKECKCKADDKTCFAEEGKGLVAASQAALGL 585
RESULT 2
US-08-702-572-2
; Sequence 2, Application US/08702572
; Patent No. 5965386
; GENERAL INFORMATION:
; APPLICANT: Kerry-Williams, Sean M
; APPLICANT: Gilbert, Sarah C
; TITLE OF INVENTION: Yeast Strains and Modified Albums
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Centeon L.L.C.
; STREET: 1020 First Avenue
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-1310
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,572
; FILING DATE: 11-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 95/23857
; FILING DATE: 1-MAR-1995
; APPLICATION NUMBER: GB 9404270.2
; FILING DATE: 5-MAR-1994

ATTORNEY/AGENT INFORMATION:
NAME: Naomi Blasas
REGISTRATION NUMBER: 38,394
REFERENCE/DOCKET NUMBER: C80114 US
TELEPHONE: 610/878/4294
TELEFAX: 610/878/4221
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-702-572-2

Query Match 100.0%; Score 585; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	DAHSEVAHRFKDGLGEENFKALVLIAPQAQYLOQCPPEHVKLVNEVTEFAKTCAVDESAB	60
DB	1	DAHSEVAHRFKDGLGEENFKALVLIAPQAQYLOQCPPEHVKLVNEVTEFAKTCAVDESAB	60
QY	61	NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEKNECFLOHKDDNPNLRLVRPEV	120
DB	61	NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEKNECFLOHKDDNPNLRLVRPEV	120
QY	121	DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP	180
DB	121	DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP	180
QY	181	KLDELDRDEGKASSAKORLKCASLOKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT	240
DB	181	KLDELDRDEGKASSAKORLKCASLOKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT	240
QY	241	VHTECCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVNDEMPE	300
DB	241	VHTECCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVNDEMPE	300
QY	301	DLPSLAADFVSKDVCCKYAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTLK	360
DB	301	DLPSLAADFVSKDVCCKYAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTLK	360
QY	361	CAAADPHECYAKVDFEFPKPLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKVPOVST	420
DB	361	CAAADPHECYAKVDFEFPKPLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKVPOVST	420
QY	421	PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVNLQNLVLTAKTYETTLK	480
DB	421	PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVNLQNLVLTAKTYETTLK	480
QY	481	LVNRRPCFSALEVDITYVPKEFNAETTFHADICTLSEKERQIKQTALVELVKKPKAT	540
DB	481	LVNRRPCFSALEVDITYVPKEFNAETTFHADICTLSEKERQIKQTALVELVKKPKAT	540
QY	541	KEQLKAVMDDFAAFVEKCKADDKETCFABEGKCLVAASQAALGL	585
DB	541	KEQLKAVMDDFAAFVEKCKADDKETCFABEGKCLVAASQAALGL	585

RESULT 3
US-08-769-746-2
Sequence 2, Application US/08769746
Patent No. 6274305
GENERAL INFORMATION:
APPLICANT: Sonnschein, Carlos
APPLICANT: Soto, Ana M.
TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco

STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/769,746
FILING DATE: 19-DEC-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: MBRI-02584
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-769-746-2

Query Match 100.0%; Score 585; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	DAHSEVAHRFKDGLGEENFKALVLIAPQAQYLOQCPPEHVKLVNEVTEFAKTCAVDESAB	60
DB	1	DAHSEVAHRFKDGLGEENFKALVLIAPQAQYLOQCPPEHVKLVNEVTEFAKTCAVDESAB	60
QY	61	NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEKNECFLOHKDDNPNLRLVRPEV	120
DB	61	NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEKNECFLOHKDDNPNLRLVRPEV	120
QY	121	DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP	180
DB	121	DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP	180
QY	181	KLDELDRDEGKASSAKORLKCASLOKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT	240
DB	181	KLDELDRDEGKASSAKORLKCASLOKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT	240
QY	241	VHTECCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVNDEMPE	300
DB	241	VHTECCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVNDEMPE	300
QY	301	DLPSLAADFVSKDVCCKYAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTLK	360
DB	301	DLPSLAADFVSKDVCCKYAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTLK	360
QY	361	CAAADPHECYAKVDFEFPKPLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKVPOVST	420
DB	361	CAAADPHECYAKVDFEFPKPLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKVPOVST	420
QY	421	PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVNLQNLVLTAKTYETTLK	480
DB	421	PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVNLQNLVLTAKTYETTLK	480
QY	481	LVNRRPCFSALEVDITYVPKEFNAETTFHADICTLSEKERQIKQTALVELVKKPKAT	540
DB	481	LVNRRPCFSALEVDITYVPKEFNAETTFHADICTLSEKERQIKQTALVELVKKPKAT	540
QY	541	KEQLKAVMDDFAAFVEKCKADDKETCFABEGKCLVAASQAALGL	585
DB	541	KEQLKAVMDDFAAFVEKCKADDKETCFABEGKCLVAASQAALGL	585

RESULT 4

US-10-153-064-5
; Sequence 5, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5

Query Match 100.0%; Score 585; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
DB 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180

QY 181 KLDELDRDEGKASSAKQRLKCAASLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVDTLTK 240
DB 181 KLDELDRDEGKASSAKQRLKCAASLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVDTLTK 240

QY 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLEKSHGICIAEVNDEMPE 300
DB 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLEKSHGICIAEVNDEMPE 300

QY 301 DPLSLAADFVESKDVCKNVAEAKDVLGMFLYELVYARRHPDYSVLLRLAKTYETTLK 360
DB 301 DPLSLAADFVESKDVCKNVAEAKDVLGMFLYELVYARRHPDYSVLLRLAKTYETTLK 360

QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
DB 361 CAAADPHECYAKVDFEFKPLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKVPQVST 420

QY 421 PTLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLSVNLQCLVLEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLSVNLQCLVLEKTPVSDRVTKCCTES 480

QY 481 LVNRRPCFSALEVDYTVPKFNAETFTFHADICTLSEKERQIKKQATLALVELVKHKPKAT 540
DB 481 LVNRRPCFSALEVDYTVPKFNAETFTFHADICTLSEKERQIKKQATLALVELVKHKPKAT 540

QY 541 KEQLKAVMDPFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 585
DB 541 KEQLKAVMDPFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 585

RESULT 5
US-10-153-064-7
; Sequence 7, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064

US-10-153-064-7
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-7

Query Match 100.0%; Score 585; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
DB 25 DAHKEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144

QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 145 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 204

QY 181 KLDELDRDEGKASSAKQRLKCAASLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVDTLTK 240
DB 205 KLDELDRDEGKASSAKQRLKCAASLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVDTLTK 264

QY 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLEKSHGICIAEVNDEMPE 300
DB 265 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLEKSHGICIAEVNDEMPE 324

QY 301 DPLSLAADFVESKDVCKNVAEAKDVLGMFLYELVYARRHPDYSVLLRLAKTYETTLK 360
DB 325 DPLSLAADFVESKDVCKNVAEAKDVLGMFLYELVYARRHPDYSVLLRLAKTYETTLK 384

QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
DB 385 CAAADPHECYAKVDFEFKPLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKVPQVST 444

QY 421 PTLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLSVNLQCLVLEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLSVNLQCLVLEKTPVSDRVTKCCTES 504

QY 481 LVNRRPCFSALEVDYTVPKFNAETFTFHADICTLSEKERQIKKQATLALVELVKHKPKAT 540
DB 505 LVNRRPCFSALEVDYTVPKFNAETFTFHADICTLSEKERQIKKQATLALVELVKHKPKAT 564

QY 541 KEQLKAVMDPFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 585
DB 565 KEQLKAVMDPFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 609

RESULT 6
US-09-976-594-977
; Sequence 977, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 977
; LENGTH: 609

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 088957CD1
US-09-976-594-977

Query Match 100.0%; Score 585; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 0; Indels 0; Gaps 0;
Matches 585; Conservative 0; Mismatches 0

QY 1 DAHSEVAHREFKDLGEENFKALVLIAPQAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60
Db 25 DAHSEVAHREFKDLGEENFKALVLIAPQAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHDDNPNLRLVPRPV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHDDNPNLRLVPRPV 144
QY 121 DVMTAFHDNDETEFLKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 180
Db 145 DVMTAFHDNDETEFLKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 204
QY 181 KLDELDRDEGKASSAKQRLKCSLOKQGERAFKAWAVARLSQRPFKAEFAVSKLVTDLT 240
Db 205 KLDELDRDEGKASSAKQRLKCSLOKQGERAFKAWAVARLSQRPFKAEFAVSKLVTDLT 264
QY 241 VHTECHGDLLECCADDDRADLAKYICENODSISSKLKECCCKPILKSHCHCIAEVNDEMPA 300
Db 265 VHTECHGDLLECCADDDRADLAKYICENODSISSKLKECCCKPILKSHCHCIAEVNDEMPA 324
QY 301 DLPSLAADFVSKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTLEKC 360
Db 325 DLPSLAADFVSKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTLEKC 384
QY 361 CAAADPHECYAKVDFEPKPLVEEPQNLKONCELFEOLGEYKFNQALLVRYTKKVPQVST 420
Db 385 CAAADPHECYAKVDFEPKPLVEEPQNLKONCELFEOLGEYKFNQALLVRYTKKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTES 480
Db 445 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTES 504
QY 481 LVNRRPCFSALVEVDITYVPKFNATFTFHADICTLSEKROIKKOTALVELVKHKPKAT 540
Db 505 LVNRRPCFSALVEVDITYVPKFNATFTFHADICTLSEKROIKKOTALVELVKHKPKAT 564
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDDKTCFAEBGKKLVAASQAALGL 585
Db 565 KEQLKAVMDDDFAAFVEKCKCKADDDKTCFAEBGKKLVAASQAALGL 609

RESULT 7
US-08-797-689-2
; Sequence 2, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleece, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guitton, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-797-689-2

Query Match 100.0%; Score 585; DB 2; Length 610;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHREFKDLGEENFKALVLIAPQAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60
Db 25 DAHSEVAHREFKDLGEENFKALVLIAPQAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHDDNPNLRLVPRPV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHDDNPNLRLVPRPV 144
QY 121 DVMTAFHDNDETEFLKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 180
Db 145 DVMTAFHDNDETEFLKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 204
QY 181 KLDELDRDEGKASSAKQRLKCSLOKQGERAFKAWAVARLSQRPFKAEFAVSKLVTDLT 240
Db 205 KLDELDRDEGKASSAKQRLKCSLOKQGERAFKAWAVARLSQRPFKAEFAVSKLVTDLT 264
QY 241 VHTECHGDLLECCADDDRADLAKYICENODSISSKLKECCCKPILKSHCHCIAEVNDEMPA 300
Db 265 VHTECHGDLLECCADDDRADLAKYICENODSISSKLKECCCKPILKSHCHCIAEVNDEMPA 324
QY 301 DLPSLAADFVSKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTLEKC 360
Db 325 DLPSLAADFVSKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTLEKC 384
QY 361 CAAADPHECYAKVDFEPKPLVEEPQNLKONCELFEOLGEYKFNQALLVRYTKKVPQVST 420
Db 385 CAAADPHECYAKVDFEPKPLVEEPQNLKONCELFEOLGEYKFNQALLVRYTKKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTES 480
Db 445 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTES 504
QY 481 LVNRRPCFSALVEVDITYVPKFNATFTFHADICTLSEKROIKKOTALVELVKHKPKAT 540
Db 505 LVNRRPCFSALVEVDITYVPKFNATFTFHADICTLSEKROIKKOTALVELVKHKPKAT 564
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDDKTCFAEBGKKLVAASQAALGL 585

Db 565 KEQLKAVMDDFAAAFVEKCKCKADDKTCFAEEGKKLVAAASQAALGL 609

RESULT 8

US-09-984-186-2
; Sequence 2, Application US/09984186
; Patent No. 686179
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittou, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3043
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,186
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808

INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-984-186-2

Query Match 100.0%; Score 585; DB 4; Length 610;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFKDLGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 25 DAHKSVAHRFKDLGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 144
QY 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFAKRYKAAFTCCQAADKAAACLLP 180

Db 145 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFAKRYKAAFTCCQAADKAAACLLP 204
QY 181 KLDELDRDEGKASSAKQRLKCSLQKFGESAFKAWAVARLSQRPFAEFAVSKLVTDLTJK 240
Db 205 KLDELDRDEGKASSAKQRLKCSLQKFGESAFKAWAVARLSQRPFAEFAVSKLVTDLTJK 264
QY 241 VHTTECHGDLLECADRADLAKYICENODSISSKLKECKEKLLEKSHCIAEVENDEMPA 300
Db 265 VHTTECHGDLLECADRADLAKYICENODSISSKLKECKEKLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYIARRHPDYSVLLLRLLAKTYETTLKRC 360
Db 325 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYIARRHPDYSVLLLRLLAKTYETTLKRC 384
QY 361 CAADPHCEYAKVDFEFPKPLVEEPQNLKQNCLEPQGEYFQNALVRYTKKVPQVST 420
Db 385 CAADPHCEYAKVDFEFPKPLVEEPQNLKQNCLEPQGEYFQNALVRYTKKVPQVST 444
QY 421 PTLVEVSRLGKVGSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRLGKVGSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDYIYVKEFNAETFTPHADICTLSEYERQIKQTALVELVKKPKAT 540
Db 505 LVNRRPCFSALEVDYIYVKEFNAETFTPHADICTLSEYERQIKQTALVELVKKPKAT 564
QY 541 KEQLKAVMDDFAAAFVEKCKCKADDKTCFAEEGKKLVAAASQAALGL 585
Db 565 KEQLKAVMDDFAAAFVEKCKCKADDKTCFAEEGKKLVAAASQAALGL 609

RESULT 9

US-10-153-064-133
; Sequence 133, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PFS56
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 133
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-133

Query Match 100.0%; Score 585; DB 4; Length 651;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFKDLGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 67 DAHKSVAHRFKDLGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 126
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 120
Db 127 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 186
QY 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFAKRYKAAFTCCQAADKAAACLLP 180
Db 187 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFAKRYKAAFTCCQAADKAAACLLP 246
QY 181 KLDELDRDEGKASSAKQRLKCSLQKFGESAFKAWAVARLSQRPFAEFAVSKLVTDLTJK 240
Db 247 KLDELDRDEGKASSAKQRLKCSLQKFGESAFKAWAVARLSQRPFAEFAVSKLVTDLTJK 306
QY 241 VHTTECHGDLLECADRADLAKYICENODSISSKLKECKEKLLEKSHCIAEVENDEMPA 300

Db 307 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCEKPLELXSHCIAEVENDEMPA 366
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEAARRHPDYSVLLLRLLAKTYETTLK 360
Db 367 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEAARRHPDYSVLLLRLLAKTYETTLK 426
Qy 361 CAAADPHECYAKVDFEFPKPLVEBPQNLIKONCELFQGLGEYFQNALVRYTKVPQVST 420
Db 427 CAAADPHECYAKVDFEFPKPLVEBPQNLIKONCELFQGLGEYFQNALVRYTKVPQVST 486
Qy 421 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 487 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 546
Qy 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATALVELVXHKPKAT 540
Db 547 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATALVELVXHKPKAT 606
Qy 541 KEQLKAVMDPFAAFVEKCKKADDKETCFABEGKULVAASQAALGL 585
Db 607 KEQLKAVMDPFAAFVEKCKKADDKETCFABEGKULVAASQAALGL 651

RESULT 10
US-10-153-064-132
; Sequence 132, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 132
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-132

Query Match 100.0%; Score 585; DB 4; Length 652;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGEEFNFKALVLIAPAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60
Db 68 DAHSEVAHRFKDLGEEFNFKALVLIAPAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 127
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRLVRPEV 120
Db 128 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRLVRPEV 187
Qy 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 188 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 247
Qy 181 KLDELDEGKASSAKORLACASLOKFGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 240
Db 248 KLDELDEGKASSAKORLACASLOKFGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 307
Qy 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCEKPLELXSHCIAEVENDEMPA 300
Db 308 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCEKPLELXSHCIAEVENDEMPA 367
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEAARRHPDYSVLLLRLLAKTYETTLK 360
Db 368 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEAARRHPDYSVLLLRLLAKTYETTLK 427
Qy 361 CAAADPHECYAKVDFEFPKPLVEBPQNLIKONCELFQGLGEYFQNALVRYTKVPQVST 420

Db 428 CAAADPHECYAKVDFEFPKPLVEBPQNLIKONCELFQGLGEYFQNALVRYTKVPQVST 487
Qy 421 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 488 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 547
Qy 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATALVELVXHKPKAT 540
Db 548 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATALVELVXHKPKAT 607
Qy 541 KEQLKAVMDPFAAFVEKCKKADDKETCFABEGKULVAASQAALGL 585
Db 608 KEQLKAVMDPFAAFVEKCKKADDKETCFABEGKULVAASQAALGL 652

RESULT 11
US-10-153-064-131
; Sequence 131, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 131
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-131

Query Match 100.0%; Score 585; DB 4; Length 653;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGEEFNFKALVLIAPAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60
Db 69 DAHSEVAHRFKDLGEEFNFKALVLIAPAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 128
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRLVRPEV 120
Db 129 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPMLPRLVRPEV 188
Qy 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 189 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 248
Qy 181 KLDELDEGKASSAKORLACASLOKFGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 240
Db 249 KLDELDEGKASSAKORLACASLOKFGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 308
Qy 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCEKPLELXSHCIAEVENDEMPA 300
Db 309 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCEKPLELXSHCIAEVENDEMPA 368
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEAARRHPDYSVLLLRLLAKTYETTLK 360
Db 369 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEAARRHPDYSVLLLRLLAKTYETTLK 428
Qy 361 CAAADPHECYAKVDFEFPKPLVEBPQNLIKONCELFQGLGEYFQNALVRYTKVPQVST 420
Db 429 CAAADPHECYAKVDFEFPKPLVEBPQNLIKONCELFQGLGEYFQNALVRYTKVPQVST 488
Qy 421 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 489 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 548
Qy 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATALVELVXHKPKAT 540

[illegible]

Query Match	100.08;	Score 585;	DB 4;	Length 656;
Best Local Similarity	100.08;	Pred. No. 0;		
Matches 585;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0
QY	1	DAHKSEVAHRFKDGLGEENFKALVLI	IAFAQYLQOCFFEDHVKLVNEVTEFAKTCVADESAE	60
Db	72	DAHKSEVAHRFKDGLGEENFKALVLI	IAFAQYLQOCFFEDHVKLVNEVTEFAKTCVADESAE	131
QY	61	NCCKSLHTLFGDKLCTVATLRET	GEWADCCAKOEPERNECSFIQHKDDNPRLVRPVP	120
Db	132	NCCKSLHTLFGDKLCTVATLRET	GEWADCCAKOEPERNECSFIQHKDDNPRLVRPVP	191
QY	121	DVMCTAFHDNBEFTFKYKLYFIAR	HPYFYAPPELLFFAKYKAAFTCCOAAKAAACLLP	180
Db	192	DVMCTAFHDNBEFTFKYKLYFIAR	HPYFYAPPELLFFAKYKAAFTCCOAAKAAACLLP	251
QY	181	KLDELRDEGKASSAKQRLKCSLOK	FGGERAFKANAVARLSQRPPEKAEFAVSKLVTDLTK	240
Db	252	KLDELRDEGKASSAKQRLKCSLOK	FGGERAFKANAVARLSQRPPEKAEFAVSKLVTDLTK	311
QY	241	VHTECHGDLLECADRADLAKYI	CENODSTSSKLKCCCKPCLKKSHCIAEVENDEMPA	300
Db	312	VHTECHGDLLECADRADLAKYI	CENODSTSSKLKCCCKPCLKKSHCIAEVENDEMPA	371
QY	301	DLPSLAADPVESKDVCKNYAKAD	VFLGMFVYEVARRHPDYSVVLRLAKTVEITTELEK	360
Db	372	DLPSLAADPVESKDVCKNYAKAD	VFLGMFVYEVARRHPDYSVVLRLAKTVEITTELEK	431
QY	361	CAAADPHECYAKVPDEFKPLVEE	PQNLIKONCELFEOQLGEYKFONALLVRYTKVQVST	420
Db	432	CAAADPHECYAKVPDEFKPLVEE	PQNLIKONCELFEOQLGEYKFONALLVRYTKVQVST	491
QY	421	PTLVEVSRNLGKVGSKCKKHPEA	KRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES	480
Db	492	PTLVEVSRNLGKVGSKCKKHPEA	KRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES	551
QY	481	LVNRRCPFSALVDEFTYVPKE	FNASTFTFHADICTLSEKEROIKKQATVVLVVKHPEKAT	540
Db	552	LVNRRCPFSALVDEFTYVPKE	FNASTFTFHADICTLSEKEROIKKQATVVLVVKHPEKAT	611
QY	541	KEOLKAVMDDDFAAFVEKCCAD	DKKETCFABEGKCLVAASQAALGL	595
Db	612	KEOLKAVMDDDFAAFVEKCCAD	DKKETCFABEGKCLVAASQAALGL	656

RESULT 13

; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 129
; LENGTH: 676
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-129

Query Match 100.0%; Score 585; DB 4; Length 676;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHFRKDLGEENFKALVLIAPAQYLOCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
Db 92 DAHSEVAHFRKDLGEENFKALVLIAPAQYLOCCPPEDHVKLVNEVTEFAKTCVADESAAE 151

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPFLRLVRPEV 120
Db 152 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPFLRLVRPEV 211

Qy 121 DVMTCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
Db 212 DVMTCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 271

Qy 181 KLDELDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
Db 272 KLDELDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 331

Qy 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 332 VHTCCGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 391

Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 360
Db 392 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 451

Qy 361 CAADPHCEYAKVDFEFPKPLVEBPQNLKQNCLEFQOLGEYKFNALLVRYTKVPQVST 420
Db 452 CAADPHCEYAKVDFEFPKPLVEBPQNLKQNCLEFQOLGEYKFNALLVRYTKVPQVST 511

Qy 421 PTLVEVSRLNKGKSGCKCKHPEAKMPCAEYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 512 PTLVEVSRLNKGKSGCKCKHPEAKMPCAEYLSVLNQLCVLHEKTPVSDRVTKCCTES 571

Qy 481 LVNRRPCFSALEVDVETVPKFEFNAETFTPHADICTLSEKEROIKQTALVELVGHKPKAT 540
Db 572 LVNRRPCFSALEVDVETVPKFEFNAETFTPHADICTLSEKEROIKQTALVELVGHKPKAT 631

Qy 541 KEOLKAVMDDFAAFEVKCKCKADKCTCFABEGKKLVAASQAALGL 585
Db 632 KEOLKAVMDDFAAFEVKCKCKADKCTCFABEGKKLVAASQAALGL 676

RESULT 15
US-10-153-064-125
; Sequence 125, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 125
; LENGTH: 677
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-123

; ORGANISM: Homo sapiens
US-10-153-064-125

Query Match 100.0%; Score 585; DB 4; Length 677;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHFRKDLGEENFKALVLIAPAQYLOCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
Db 93 DAHSEVAHFRKDLGEENFKALVLIAPAQYLOCCPPEDHVKLVNEVTEFAKTCVADESAAE 152

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPFLRLVRPEV 120
Db 153 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPFLRLVRPEV 212

Qy 121 DVMTCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
Db 213 DVMTCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 272

Qy 181 KLDELDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
Db 273 KLDELDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 332

Qy 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 333 VHTCCGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 392

Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 360
Db 393 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEYARRHPDYSVLLRLAKTYETTTLEKC 452

Qy 361 CAADPHCEYAKVDFEFPKPLVEBPQNLKQNCLEFQOLGEYKFNALLVRYTKVPQVST 420
Db 453 CAADPHCEYAKVDFEFPKPLVEBPQNLKQNCLEFQOLGEYKFNALLVRYTKVPQVST 512

Qy 421 PTLVEVSRLNKGKSGCKCKHPEAKMPCAEYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 513 PTLVEVSRLNKGKSGCKCKHPEAKMPCAEYLSVLNQLCVLHEKTPVSDRVTKCCTES 572

Qy 481 LVNRRPCFSALEVDVETVPKFEFNAETFTPHADICTLSEKEROIKQTALVELVGHKPKAT 540
Db 573 LVNRRPCFSALEVDVETVPKFEFNAETFTPHADICTLSEKEROIKQTALVELVGHKPKAT 632

Qy 541 KEOLKAVMDDFAAFEVKCKCKADKCTCFABEGKKLVAASQAALGL 585
Db 633 KEOLKAVMDDFAAFEVKCKCKADKCTCFABEGKKLVAASQAALGL 677

RESULT 16
US-10-153-064-123
; Sequence 123, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 123
; LENGTH: 680
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-123

Query Match 100.0%; Score 585; DB 4; Length 680;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHFRKDLGEENFKALVLIAPAQYLOCCPPEDHVKLVNEVTEFAKTCVADESAAE 60

Db 96 DAHSEVAHRFKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNEVTEPAKTCVADESAE 155
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVVRPV 120
Db 156 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVVRPV 215
QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 180
Db 216 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 275
QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFAKAWAVARLSQRPKABFAEVSCLVTDLT 240
Db 276 KLDELDEGKASSAKQRLKCSLQKFGERAFAKAWAVARLSQRPKABFAEVSCLVTDLT 335
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPFLLEKSHCIAEVENDEMPA 300
Db 336 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPFLLEKSHCIAEVENDEMPA 395
QY 301 DLPISLAADFVESKDVCKNYAEAKVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLK 360
Db 396 DLPISLAADFVESKDVCKNYAEAKVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLK 455
QY 361 CAADDPHECYAKVDFEPKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420
Db 456 CAADDPHECYAKVDFEPKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 515
QY 421 PTLVEVSRLNGLKVGSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 516 PTLVEVSRLNGLKVGSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 575
QY 481 LVNRRPCFSALEVDVETVYKPFNAETFFHADICTLSEKERQIKKOTALVELVGHKPKAT 540
Db 576 LVNRRPCFSALEVDVETVYKPFNAETFFHADICTLSEKERQIKKOTALVELVGHKPKAT 635
QY 541 KEQLKAVMDDFAAFEVKCKADDDKTCFAEKGKLVAAASQAALGL 585
Db 636 KEQLKAVMDDFAAFEVKCKADDDKTCFAEKGKLVAAASQAALGL 680

RESULT 17
US-08-256-938-2
; Sequence 2, Application US/08256938
; Patent No. 5665863
; GENERAL INFORMATION:
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE
; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.0 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,938
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/01065
; FILING DATE: 31-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Rosanne
; REGISTRATION NUMBER: 32,534
; REFERENCE/DOCKET NUMBER: ST32007-US

TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3817
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 783 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-256-938-2

Query Match 100.0%; Score 585; DB 1; Length 783;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNEVTEPAKTCVADESAE 60
Db 25 DAHSEVAHRFKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNEVTEPAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVVRPV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVVRPV 144
QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 180
Db 145 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 204
QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFAKAWAVARLSQRPKABFAEVSCLVTDLT 240
Db 205 KLDELDEGKASSAKQRLKCSLQKFGERAFAKAWAVARLSQRPKABFAEVSCLVTDLT 264
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPFLLEKSHCIAEVENDEMPA 300
Db 265 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPFLLEKSHCIAEVENDEMPA 324
QY 301 DLPISLAADFVESKDVCKNYAEAKVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLK 360
Db 325 DLPISLAADFVESKDVCKNYAEAKVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLK 384
QY 361 CAADDPHECYAKVDFEPKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420
Db 385 CAADDPHECYAKVDFEPKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRLNGLKVGSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRLNGLKVGSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDVETVYKPFNAETFFHADICTLSEKERQIKKOTALVELVGHKPKAT 540
Db 505 LVNRRPCFSALEVDVETVYKPFNAETFFHADICTLSEKERQIKKOTALVELVGHKPKAT 564
QY 541 KEQLKAVMDDFAAFEVKCKADDDKTCFAEKGKLVAAASQAALGL 585
Db 565 KEQLKAVMDDFAAFEVKCKADDDKTCFAEKGKLVAAASQAALGL 609

RESULT 18
US-08-256-938-4
; Sequence 4, Application US/08256938
; Patent No. 5665863
; GENERAL INFORMATION:
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE
; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.0 (PatentIn)
CURRENT APPLICATION DATA: US/08/256,938
FILING DATE: 31-JAN-1992
CLASSIFICATION: 435
PRIORITY APPLICATION NUMBER: FR 92/01065
FILING DATE: 31-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Goodman, Rosanne
REGISTRATION NUMBER: 32,534
REFERENCE/DOCKET NUMBER: ST92007-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3817
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 787 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-256-938-4

Query Match 100.0%; Score 585; DB 1; Length 787;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGSENFALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 203 DAHSEVAHFRKDLGSENFALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 262

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDDNPRLVLRPEV 120
Db 263 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDDNPRLVLRPEV 322

QY 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFAPPELLFFAKYKAAFTCCQAADKAACLLP 180
Db 323 DVMTAFHDNEETFLKKYLYEIAARRHPYFAPPELLFFAKYKAAFTCCQAADKAACLLP 382

QY 181 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
Db 383 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 442

QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300

Query Match 100.0%; Score 585; DB 2; Length 787;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGSENFALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 203 DAHSEVAHFRKDLGSENFALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 262

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDDNPRLVLRPEV 120
Db 263 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDDNPRLVLRPEV 322

QY 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFAPPELLFFAKYKAAFTCCQAADKAACLLP 180
Db 323 DVMTAFHDNEETFLKKYLYEIAARRHPYFAPPELLFFAKYKAAFTCCQAADKAACLLP 382

QY 181 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
Db 383 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 442

QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300

Query Match 100.0%; Score 585; DB 2; Length 787;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGSENFALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 203 DAHSEVAHFRKDLGSENFALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 262

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDDNPRLVLRPEV 120
Db 263 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDDNPRLVLRPEV 322

QY 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFAPPELLFFAKYKAAFTCCQAADKAACLLP 180
Db 323 DVMTAFHDNEETFLKKYLYEIAARRHPYFAPPELLFFAKYKAAFTCCQAADKAACLLP 382

QY 181 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
Db 383 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 442

QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300

Query Match 100.0%; Score 585; DB 2; Length 787;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGSENFALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 203 DAHSEVAHFRKDLGSENFALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 262

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDDNPRLVLRPEV 120
Db 263 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDDNPRLVLRPEV 322

QY 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFAPPELLFFAKYKAAFTCCQAADKAACLLP 180
Db 323 DVMTAFHDNEETFLKKYLYEIAARRHPYFAPPELLFFAKYKAAFTCCQAADKAACLLP 382

QY 181 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
Db 383 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 442

QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300

Db 443 VHTCCGDLLECADRADLAKYICENQDSISSKKECKEPLLEKSHCIAEVENDMEPA 502
QY 301 DLPSLAADFVESKOVCKYAKADVFLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
Db 503 DLPSLAADFVESKOVCKYAKADVFLGMFLYEVARRHPDYSVLLRLAKTYETTLK 562
QY 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKONCELFQELGEYKFNQALLVRYTKKVPQVST 420
Db 563 CAADPHCEYAKVDFEFKPLVEEPQNLIKONCELFQELGEYKFNQALLVRYTKKVPQVST 622
QY 421 PTLVEVSRLKGVSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 623 PTLVEVSRLKGVSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 682
QY 481 LVNRRPCFSALEVDETVYPKFNATFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 540
Db 683 LVNRRPCFSALEVDETVYPKFNATFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 742
QY 541 KEQLKAVNDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
Db 743 KEQLKAVNDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 787

RESULT 20
US-09-984-186-16
; Sequence 16, Application US/09984186
; Patent No. 6686179
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Gerard
; Yen, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/984,186
FILING DATE: 29-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 787 amino acids

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-984-186-16
Query Match 100.0%; Score 585; DB 4; Length 787;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHFKDLGRENFKALVLIAPAYLOQCPEPDHVKLVNEVTBFAKTCVADESAE 60
Db 203 DAHSEVAHFKDLGRENFKALVLIAPAYLOQCPEPDHVKLVNEVTBFAKTCVADESAE 262
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLOHKDDNPNLPRLVREP 120
Db 263 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLOHKDDNPNLPRLVREP 322
QY 121 DVMCTAFHDNEETFLKKLYEIARRHPYFAPPELLFFAKRYKAAFTCCQADKAAACLLP 180
Db 323 DVMCTAFHDNEETFLKKLYEIARRHPYFAPPELLFFAKRYKAAFTCCQADKAAACLLP 382
QY 181 KLDELDRDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRFPKAEPAEVSKLVTDLTK 240
Db 383 KLDELDRDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRFPKAEPAEVSKLVTDLTK 442
QY 241 VHTCCGDLLECADRADLAKYICENQDSISSKKECKEPLLEKSHCIAEVENDMEPA 300
Db 443 VHTCCGDLLECADRADLAKYICENQDSISSKKECKEPLLEKSHCIAEVENDMEPA 502
QY 301 DLPSLAADFVESKOVCKYAKADVFLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
Db 503 DLPSLAADFVESKOVCKYAKADVFLGMFLYEVARRHPDYSVLLRLAKTYETTLK 562
QY 361 CAADPHCEYAKVDFEFKPLVEEPQNLIKONCELFQELGEYKFNQALLVRYTKKVPQVST 420
Db 563 CAADPHCEYAKVDFEFKPLVEEPQNLIKONCELFQELGEYKFNQALLVRYTKKVPQVST 622
QY 421 PTLVEVSRLKGVSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 623 PTLVEVSRLKGVSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 682
QY 481 LVNRRPCFSALEVDETVYPKFNATFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 540
Db 683 LVNRRPCFSALEVDETVYPKFNATFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 742
QY 541 KEQLKAVNDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
Db 743 KEQLKAVNDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 787

RESULT 21
US-10-153-064-96
; Sequence 96, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 96
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-96
Query Match 87.2%; Score 510; DB 4; Length 652;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGSENFKALVLIAPQYLOQCPEFEDHVKLVNEVTEFAKTCVADESAAE 60
 Db 67 DAHSEVAHRFKDLGSENFKALVLIAPQYLOQCPEFEDHVKLVNEVTEFAKTCVADESAAE 126

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVPRPEV 120
 Db 127 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVPRPEV 186

Qy 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
 Db 187 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 246

Qy 181 KLDELDRGKASSAKORLKASLOKFGERAFAKAWARLSORFPKAEVSKLVTDLTK 240
 Db 247 KLDELDRGKASSAKORLKASLOKFGERAFAKAWARLSORFPKAEVSKLVTDLTK 306

Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLEKSHCIAEVENDEMPA 300
 Db 307 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLEKSHCIAEVENDEMPA 366

Qy 361 CAAADPHCEYAKVDFEFPKPLVEEPQNLIKONCELPQGLGEYKFNALLVRYTKKVPQVST 420
 Db 427 CAAADPHCEYAKVDFEFPKPLVEEPQNLIKONCELPQGLGEYKFNALLVRYTKKVPQVST 486

Qy 421 PTLVEVRNLGKVGSKCKKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 Db 487 PTLVEVRNLGKVGSKCKKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 546

Qy 481 LVNRRPCFSALEVDVETVVPKFNATFTFH 510
 Db 547 LVNRRPCFSALEVDVETVVPKFNATFTFH 576

RESULT 22

US-10-153-064-99

; Sequence 99, Application US/10153064

; Patent No. 6663485

; GENERAL INFORMATION:

; APPLICANT: Bell et al.

; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins

; FILE REFERENCE: PF556

; CURRENT APPLICATION NUMBER: US/10/153,064

; CURRENT FILING DATE: 2002-05-24

; PRIOR APPLICATION NUMBER: 60/293,212

; PRIOR FILING DATE: 2001-05-25

; NUMBER OF SEQ ID NOS: 137

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 99

; LENGTH: 652

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-153-064-99

Query Match 87.2%; Score 510; DB 4; Length 652;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGSENFKALVLIAPQYLOQCPEFEDHVKLVNEVTEFAKTCVADESAAE 60
 Db 67 DAHSEVAHRFKDLGSENFKALVLIAPQYLOQCPEFEDHVKLVNEVTEFAKTCVADESAAE 126

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVPRPEV 120
 Db 127 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVPRPEV 186

Qy 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180

Db 187 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 246

Qy 181 KLDELDRGKASSAKORLKASLOKFGERAFAKAWARLSORFPKAEVSKLVTDLTK 240
 Db 247 KLDELDRGKASSAKORLKASLOKFGERAFAKAWARLSORFPKAEVSKLVTDLTK 306

Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLEKSHCIAEVENDEMPA 300
 Db 307 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLEKSHCIAEVENDEMPA 366

Qy 361 CAAADPHCEYAKVDFEFPKPLVEEPQNLIKONCELPQGLGEYKFNALLVRYTKKVPQVST 420
 Db 427 CAAADPHCEYAKVDFEFPKPLVEEPQNLIKONCELPQGLGEYKFNALLVRYTKKVPQVST 486

Qy 421 PTLVEVRNLGKVGSKCKKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 Db 487 PTLVEVRNLGKVGSKCKKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 546

Qy 481 LVNRRPCFSALEVDVETVVPKFNATFTFH 510
 Db 547 LVNRRPCFSALEVDVETVVPKFNATFTFH 576

RESULT 23

US-10-153-064-105

; Sequence 105, Application US/10153064

; Patent No. 6663485

; GENERAL INFORMATION:

; APPLICANT: Bell et al.

; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins

; FILE REFERENCE: PF556

; CURRENT APPLICATION NUMBER: US/10/153,064

; CURRENT FILING DATE: 2002-05-24

; PRIOR APPLICATION NUMBER: 60/293,212

; PRIOR FILING DATE: 2001-05-25

; NUMBER OF SEQ ID NOS: 137

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 105

; LENGTH: 652

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-153-064-105

Query Match 87.2%; Score 510; DB 4; Length 652;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGSENFKALVLIAPQYLOQCPEFEDHVKLVNEVTEFAKTCVADESAAE 60
 Db 67 DAHSEVAHRFKDLGSENFKALVLIAPQYLOQCPEFEDHVKLVNEVTEFAKTCVADESAAE 126

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVPRPEV 120
 Db 127 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVPRPEV 186

Qy 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
 Db 187 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 246

Qy 181 KLDELDRGKASSAKORLKASLOKFGERAFAKAWARLSORFPKAEVSKLVTDLTK 240
 Db 247 KLDELDRGKASSAKORLKASLOKFGERAFAKAWARLSORFPKAEVSKLVTDLTK 306

Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLEKSHCIAEVENDEMPA 300
 Db 307 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLEKSHCIAEVENDEMPA 366

Qy 361 CAAADPHCEYAKVDFEFPKPLVEEPQNLIKONCELPQGLGEYKFNALLVRYTKKVPQVST 420
 Db 427 CAAADPHCEYAKVDFEFPKPLVEEPQNLIKONCELPQGLGEYKFNALLVRYTKKVPQVST 486

Db 367 DLPSLAADFVSKDVKYAEAKDVLGMFLYFYARRHPDYSVVLRLAKTYETTLK 426
QY 361 CAADPHCEYAKVDFEPKPLVEEPQNLKONCELFQOLGEYKFNALLVRYTKVPQVST 420
Db 427 CAADPHCEYAKVDFEPKPLVEEPQNLKONCELFQOLGEYKFNALLVRYTKVPQVST 486
QY 421 PTLVEVSRNLGKVGSKCKHPKAEKMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 487 PTLVEVSRNLGKVGSKCKHPKAEKMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 546
QY 481 LVNRRPCFSALEVDVETVPKEFNAETTFH 510
Db 547 LVNRRPCFSALEVDVETVPKEFNAETTFH 576

RESULT 24

US-10-153-064-90
; Sequence 90, Application US/10153064
; Patent No. 6663485

GENERAL INFORMATION:

; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins

; FILE REFERENCE: PF556

; CURRENT APPLICATION NUMBER: US/10/153,064

; CURRENT FILING DATE: 2002-05-24

; PRIOR APPLICATION NUMBER: 60/293,212

; PRIOR FILING DATE: 2001-05-25

; NUMBER OF SEQ ID NOS: 137

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 90

; LENGTH: 660

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-153-064-90

Query Match 87.2%; Score 510; DB 4; Length 660;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDGLGEENFKALVLIAPQYLOQCPFDHVKLVNEVTEFAKTCVADESAB 60
Db 75 DAHSEVAHRFKDGLGEENFKALVLIAPQYLOQCPFDHVKLVNEVTEFAKTCVADESAB 134

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPPLRLVRPEV 120
Db 135 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPPLRLVRPEV 194

QY 121 DVMTAFHDNEETFLKKYLYEIAERHPYFYAPPELLFPKGYKAAFTCCQADKAACLLP 180
Db 195 DVMTAFHDNEETFLKKYLYEIAERHPYFYAPPELLFPKGYKAAFTCCQADKAACLLP 254

QY 181 KLDELDRDEGKASSAKQRLKCAQLQKFGERAFKAWAVARLSQRPKAEFAVSKLVTDLT 240
Db 255 KLDELDRDEGKASSAKQRLKCAQLQKFGERAFKAWAVARLSQRPKAEFAVSKLVTDLT 314

QY 241 VHTCCGHDLLCEADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 315 VHTCCGHDLLCEADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 374

QY 301 DLPSLAADFVSKDVKYAEAKDVLGMFLYFYARRHPDYSVVLRLAKTYETTLK 360
Db 375 DLPSLAADFVSKDVKYAEAKDVLGMFLYFYARRHPDYSVVLRLAKTYETTLK 434

QY 361 CAADPHCEYAKVDFEPKPLVEEPQNLKONCELFQOLGEYKFNALLVRYTKVPQVST 420
Db 435 CAADPHCEYAKVDFEPKPLVEEPQNLKONCELFQOLGEYKFNALLVRYTKVPQVST 494

QY 421 PTLVEVSRNLGKVGSKCKHPKAEKMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 495 PTLVEVSRNLGKVGSKCKHPKAEKMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 554

QY 481 LVNRRPCFSALEVDVETVPKEFNAETTFH 510
Db 547 LVNRRPCFSALEVDVETVPKEFNAETTFH 576

Db 555 LVNRRPCFSALEVDVETVPKEFNAETTFH 584

RESULT 25

US-10-153-064-93

; Sequence 93, Application US/10153064

; Patent No. 6663485

GENERAL INFORMATION:

; APPLICANT: Bell et al.

; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins

; FILE REFERENCE: PF556

; CURRENT APPLICATION NUMBER: US/10/153,064

; CURRENT FILING DATE: 2002-05-24

; PRIOR APPLICATION NUMBER: 60/293,212

; PRIOR FILING DATE: 2001-05-25

; NUMBER OF SEQ ID NOS: 137

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 93

; LENGTH: 660

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-153-064-93

Query Match 87.2%; Score 510; DB 4; Length 660;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDGLGEENFKALVLIAPQYLOQCPFDHVKLVNEVTEFAKTCVADESAB 60
Db 75 DAHSEVAHRFKDGLGEENFKALVLIAPQYLOQCPFDHVKLVNEVTEFAKTCVADESAB 134

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPPLRLVRPEV 120
Db 135 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPPLRLVRPEV 194

QY 121 DVMTAFHDNEETFLKKYLYEIAERHPYFYAPPELLFPKGYKAAFTCCQADKAACLLP 180
Db 195 DVMTAFHDNEETFLKKYLYEIAERHPYFYAPPELLFPKGYKAAFTCCQADKAACLLP 254

QY 181 KLDELDRDEGKASSAKQRLKCAQLQKFGERAFKAWAVARLSQRPKAEFAVSKLVTDLT 240
Db 255 KLDELDRDEGKASSAKQRLKCAQLQKFGERAFKAWAVARLSQRPKAEFAVSKLVTDLT 314

QY 241 VHTCCGHDLLCEADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 315 VHTCCGHDLLCEADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 374

QY 301 DLPSLAADFVSKDVKYAEAKDVLGMFLYFYARRHPDYSVVLRLAKTYETTLK 360
Db 375 DLPSLAADFVSKDVKYAEAKDVLGMFLYFYARRHPDYSVVLRLAKTYETTLK 434

QY 361 CAADPHCEYAKVDFEPKPLVEEPQNLKONCELFQOLGEYKFNALLVRYTKVPQVST 420
Db 435 CAADPHCEYAKVDFEPKPLVEEPQNLKONCELFQOLGEYKFNALLVRYTKVPQVST 494

QY 421 PTLVEVSRNLGKVGSKCKHPKAEKMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 495 PTLVEVSRNLGKVGSKCKHPKAEKMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 554

QY 481 LVNRRPCFSALEVDVETVPKEFNAETTFH 510
Db 555 LVNRRPCFSALEVDVETVPKEFNAETTFH 584

RESULT 26

US-10-153-064-95

; Sequence 95, Application US/10153064

; Patent No. 6663485

GENERAL INFORMATION:

; APPLICANT: Bell et al.

; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins

; FILE REFERENCE: PF556

; CURRENT APPLICATION NUMBER: US/10/153,064

```
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 95
; LENGTH: 676
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-95

Query Match      87.2%; Score 510; DB 4; Length 676;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDGLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 60
Db 91 DAHSEVAHRFKDGLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 150

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLOHKDNDPNLPRLVPEV 120
Db 151 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLOHKDNDPNLPRLVPEV 210

Qy 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 211 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 270

Qy 181 KLDELDEGKASSAKORLKASLQKFGERAFAKAWARLSORFPKAEFAEYSKLVTDLT 240
Db 271 KLDELDEGKASSAKORLKASLQKFGERAFAKAWARLSORFPKAEFAEYSKLVTDLT 330

Qy 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 331 VHTTECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 390

Qy 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEYARRHPDYSVLLLRLLAKTYETTTLEK 360
Db 391 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEYARRHPDYSVLLLRLLAKTYETTTLEK 450

Qy 361 CAAADPHECYAKVDFEFPKPLVEEPQNLKONCELFQNGELGEYKFNALLVRYTKVPQVST 420
Db 451 CAAADPHECYAKVDFEFPKPLVEEPQNLKONCELFQNGELGEYKFNALLVRYTKVPQVST 510

Qy 421 PTLVEVSRNLGKVGSKCKHPKAKMPCAEADYLSVNLQCLVHKEKTPVSDRVTKCCTES 480
Db 511 PTLVEVSRNLGKVGSKCKHPKAKMPCAEADYLSVNLQCLVHKEKTPVSDRVTKCCTES 570

Qy 481 LVNRRPCFSALEVDYTVPKFNAETFTFH 510
Db 571 LVNRRPCFSALEVDYTVPKFNAETFTFH 600

RESULT 27
US-10-153-064-98
; Sequence 98, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 98
; LENGTH: 676
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-98

Query Match      87.2%; Score 510; DB 4; Length 676;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDGLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 60
Db 91 DAHSEVAHRFKDGLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 150

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLOHKDNDPNLPRLVPEV 120
Db 151 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLOHKDNDPNLPRLVPEV 210

Qy 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 211 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 270

Qy 181 KLDELDEGKASSAKORLKASLQKFGERAFAKAWARLSORFPKAEFAEYSKLVTDLT 240
Db 271 KLDELDEGKASSAKORLKASLQKFGERAFAKAWARLSORFPKAEFAEYSKLVTDLT 330

Qy 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 331 VHTTECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 390

Qy 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEYARRHPDYSVLLLRLLAKTYETTTLEK 360
Db 391 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEYARRHPDYSVLLLRLLAKTYETTTLEK 450

Qy 361 CAAADPHECYAKVDFEFPKPLVEEPQNLKONCELFQNGELGEYKFNALLVRYTKVPQVST 420
Db 451 CAAADPHECYAKVDFEFPKPLVEEPQNLKONCELFQNGELGEYKFNALLVRYTKVPQVST 510

Qy 421 PTLVEVSRNLGKVGSKCKHPKAKMPCAEADYLSVNLQCLVHKEKTPVSDRVTKCCTES 480
Db 511 PTLVEVSRNLGKVGSKCKHPKAKMPCAEADYLSVNLQCLVHKEKTPVSDRVTKCCTES 570

Qy 481 LVNRRPCFSALEVDYTVPKFNAETFTFH 510
Db 571 LVNRRPCFSALEVDYTVPKFNAETFTFH 600
```

```
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDGLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 60
Db 91 DAHSEVAHRFKDGLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 150

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLOHKDNDPNLPRLVPEV 120
Db 151 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLOHKDNDPNLPRLVPEV 210

Qy 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 211 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 270

Qy 181 KLDELDEGKASSAKORLKASLQKFGERAFAKAWARLSORFPKAEFAEYSKLVTDLT 240
Db 271 KLDELDEGKASSAKORLKASLQKFGERAFAKAWARLSORFPKAEFAEYSKLVTDLT 330

Qy 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 331 VHTTECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 390

Qy 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEYARRHPDYSVLLLRLLAKTYETTTLEK 360
Db 391 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEYARRHPDYSVLLLRLLAKTYETTTLEK 450

Qy 361 CAAADPHECYAKVDFEFPKPLVEEPQNLKONCELFQNGELGEYKFNALLVRYTKVPQVST 420
Db 451 CAAADPHECYAKVDFEFPKPLVEEPQNLKONCELFQNGELGEYKFNALLVRYTKVPQVST 510

Qy 421 PTLVEVSRNLGKVGSKCKHPKAKMPCAEADYLSVNLQCLVHKEKTPVSDRVTKCCTES 480
Db 511 PTLVEVSRNLGKVGSKCKHPKAKMPCAEADYLSVNLQCLVHKEKTPVSDRVTKCCTES 570

Qy 481 LVNRRPCFSALEVDYTVPKFNAETFTFH 510
Db 571 LVNRRPCFSALEVDYTVPKFNAETFTFH 600

RESULT 28
US-10-153-064-104
; Sequence 104, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 104
; LENGTH: 676
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-104

Query Match      87.2%; Score 510; DB 4; Length 676;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDGLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 60
Db 91 DAHSEVAHRFKDGLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 150

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLOHKDNDPNLPRLVPEV 120
Db 151 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLOHKDNDPNLPRLVPEV 210

Qy 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180
```

Db 211 DVMTAFHNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 270
Qy 181 KLDELDEGKASSAKQRLKASLOKGERAFKAWAVARLSQPPKAPFAEVSCLVTDLTk 240
Db 271 KLDELDEGKASSAKQRLKASLOKGERAFKAWAVARLSQPPKAPFAEVSCLVTDLTk 330
Qy 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLECCCKPILLESKSHCIAEVENDEMPA 300
Db 331 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLECCCKPILLESKSHCIAEVENDEMPA 390
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIARRHPDYSVLLLRLLAKTYETTTLEK 360
Db 391 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIARRHPDYSVLLLRLLAKTYETTTLEK 450
Qy 361 CAADPHCEYAKVDFEKLVEBPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420
Db 451 CAADPHCEYAKVDFEKLVEBPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 510
Qy 421 PTLVEVRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 511 PTLVEVRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 570
Qy 481 LVNRRPCFSALEVDETVVPKEFNAETFTFH 510
Db 571 LVNRRPCFSALEVDETVVPKEFNAETFTFH 600

RESULT 29

US-10-153-064-92
; Sequence 92, Application US/10153064
; Patent No. 6663485

; GENERAL INFORMATION:
; APPLICANT: Bell et al.

; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556

; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24

; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25

; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patent in version 3.1

; SEQ ID NO 92
; LENGTH: 684

; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-153-064-92

Query Match 87.2%; Score 510; DB 4; Length 684;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60
Db 99 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 158
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLPRLVREV 120
Db 159 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLPRLVREV 218
Qy 121 DVMTAFHNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 180
Db 219 DVMTAFHNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 278
Qy 181 KLDELDEGKASSAKQRLKASLOKGERAFKAWAVARLSQPPKAPFAEVSCLVTDLTk 240
Db 279 KLDELDEGKASSAKQRLKASLOKGERAFKAWAVARLSQPPKAPFAEVSCLVTDLTk 338
Qy 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLECCCKPILLESKSHCIAEVENDEMPA 300
Db 339 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLECCCKPILLESKSHCIAEVENDEMPA 398
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIARRHPDYSVLLLRLLAKTYETTTLEK 360

Db 399 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIARRHPDYSVLLLRLLAKTYETTTLEK 458
Qy 361 CAADPHCEYAKVDFEKLVEBPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420
Db 459 CAADPHCEYAKVDFEKLVEBPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 518
Qy 421 PTLVEVRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 519 PTLVEVRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 578
Qy 481 LVNRRPCFSALEVDETVVPKEFNAETFTFH 510
Db 579 LVNRRPCFSALEVDETVVPKEFNAETFTFH 608

RESULT 30

US-10-153-064-89
; Sequence 89, Application US/10153064
; Patent No. 6663485

; GENERAL INFORMATION:
; APPLICANT: Bell et al.

; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556

; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24

; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25

; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patent in version 3.1

; SEQ ID NO 89
; LENGTH: 1184

; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-153-064-89

Query Match 87.2%; Score 510; DB 4; Length 1184;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60
Db 599 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 658
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLPRLVREV 120
Db 559 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLPRLVREV 718
Qy 121 DVMTAFHNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 180
Db 719 DVMTAFHNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 778
Qy 181 KLDELDEGKASSAKQRLKASLOKGERAFKAWAVARLSQPPKAPFAEVSCLVTDLTk 240
Db 779 KLDELDEGKASSAKQRLKASLOKGERAFKAWAVARLSQPPKAPFAEVSCLVTDLTk 838
Qy 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLECCCKPILLESKSHCIAEVENDEMPA 300
Db 839 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKLECCCKPILLESKSHCIAEVENDEMPA 898
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIARRHPDYSVLLLRLLAKTYETTTLEK 360
Db 899 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIARRHPDYSVLLLRLLAKTYETTTLEK 958
Qy 361 CAADPHCEYAKVDFEKLVEBPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420
Db 959 CAADPHCEYAKVDFEKLVEBPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 1018
Qy 421 PTLVEVRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 1019 PTLVEVRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 1078
Qy 481 LVNRRPCFSALEVDETVVPKEFNAETFTFH 510

Db 1079 LVNRRCPFSALEVDYVVPKFNATFTFH 1108
RESULT 31
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri
; APPLICANT: Wurfel, David
; APPLICANT: Lyons, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Angen Center, Patent Operations/REC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,619
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-222-619-3

Query Match 82.7%; Score 484; DB 1; Length 609;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 584; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 DAHSEVAHREKDLGEENFKALVLIIFAQYLOQCPPEHVKLVNEVTEFAKTCVADESAE 60
DB 25 DAHSEVAHREKDLGEENFKALVLIIFAQYLOQCPPEHVKLVNEVTEFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVMTCTAFHDNEETFLKYLVEIARRHPYFVAPELLFPKRYKAAFTCCOQADKAAACLLP 180
DB 145 DVMTCTAFHDNEETFLKYLVEIARRHPYFVAPELLFPKRYKAAFTCCOQADKAAACLLP 204
QY 181 KLDELDRDEGKASSAKORLKCSLQKGERAFKAWAVARLSQRFPAEFVSKLVTDLTK 240
DB 205 KLDELDRDEGKASSAKORLKCSLQKGERAFKAWAVARLSQRFPAEFVSKLVTDLTK 264
QY 241 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLECCCKPILKSHCHIAEVNDEMA 300
DB 265 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLECCCKPILKSHCHIAEVNDEMA 324
QY 301 DLPSLAADFVSKDKCKNYAEAKDVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLKRC 360
DB 325 DLPSLAADFVSKDKCKNYAEAKDVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLKRC 384
QY 361 CAAADPHECYAKVDFEFLPVEEONLIKONCELFQGLGEYKQFONALLVRYTKVQVST 420
DB 385 CAAADPHECYAKVDFEFLPVEEONLIKONCELFQGLGEYKQFONALLVRYTKVQVST 444

Query Match 82.7%; Score 484; DB 1; Length 609;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 584; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 DAHSEVAHREKDLGEENFKALVLIIFAQYLOQCPPEHVKLVNEVTEFAKTCVADESAE 60
DB 25 DAHSEVAHREKDLGEENFKALVLIIFAQYLOQCPPEHVKLVNEVTEFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVMTCTAFHDNEETFLKYLVEIARRHPYFVAPELLFPKRYKAAFTCCOQADKAAACLLP 180
DB 145 DVMTCTAFHDNEETFLKYLVEIARRHPYFVAPELLFPKRYKAAFTCCOQADKAAACLLP 204
QY 181 KLDELDRDEGKASSAKORLKCSLQKGERAFKAWAVARLSQRFPAEFVSKLVTDLTK 240
DB 205 KLDELDRDEGKASSAKORLKCSLQKGERAFKAWAVARLSQRFPAEFVSKLVTDLTK 264
QY 241 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLECCCKPILKSHCHIAEVNDEMA 300
DB 265 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLECCCKPILKSHCHIAEVNDEMA 324
QY 301 DLPSLAADFVSKDKCKNYAEAKDVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLKRC 360
DB 325 DLPSLAADFVSKDKCKNYAEAKDVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLKRC 384
QY 361 CAAADPHECYAKVDFEFLPVEEONLIKONCELFQGLGEYKQFONALLVRYTKVQVST 420
DB 385 CAAADPHECYAKVDFEFLPVEEONLIKONCELFQGLGEYKQFONALLVRYTKVQVST 444

RESULT 32
US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828
; GENERAL INFORMATION:
; APPLICANT: Steekrishna, Kotikanyadan
; APPLICANT: Barr, Kathryn A.
; APPLICANT: Brierley, Russell A.
; APPLICANT: Thill, Gregory P.
; APPLICANT: Tschoep, Juerg F.
; TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN
; TITLE OF INVENTION: PICHIA PASTORIS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/433,037
; FILING DATE: 03-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9108Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-433-037-4

Query Match 82.7%; Score 484; DB 1; Length 609;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 584; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 DAHSEVAHREKDLGEENFKALVLIIFAQYLOQCPPEHVKLVNEVTEFAKTCVADESAE 60
DB 25 DAHSEVAHREKDLGEENFKALVLIIFAQYLOQCPPEHVKLVNEVTEFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVMTCTAFHDNEETFLKYLVEIARRHPYFVAPELLFPKRYKAAFTCCOQADKAAACLLP 180
DB 145 DVMTCTAFHDNEETFLKYLVEIARRHPYFVAPELLFPKRYKAAFTCCOQADKAAACLLP 204

Db 265 VHTCCGGLLESCADDRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 324
QY 301 DPLSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYSVLLLRLLAKTYETTLTKC 360
Db 325 DPLSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYSVLLLRLLAKTYETTLTKC 384
QY 361 CAADPHCEYAKVDFEFPKPLVEEPQNLKONCELFEQLGEYKFNALLVRYTKVQPVYST 420
Db 385 CAADPHCEYAKVDFEFPKPLVEEPQNLKONCELFEQLGEYKFNALLVRYTKVQPVYST 444
QY 421 PTLVEVSRNLGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVEDTVYPKEFNAETFFHADICTLSEKEROIKKOTALVELVKHKPKAT 540
Db 505 LVNRRPCFSALEVEDTVYPKEFNAETFFHADICTLSEKEROIKKOTALVELVKHKPKAT 564
QY 541 KEQLKAVMDDFAAVFKECKCKADDKETCFABEGKLVAAASQAALGL 585
Db 565 KEQLKAVMDDFAAVFKECKCKADDKETCFABEGKLVAAASQAALGL 609

RESULT 34

PCT-US95-04075-3
; Sequence 3, Application PC/TUS9504075
; GENERAL INFORMATION:
; APPLICANT: AMGEN INC.
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04075
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; PCT-US95-04075-3

Query Match 82.7%; Score 484; DB 5; Length 609;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 584; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
Db 25 DAHSEVAHRFKDLGGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 84
QY 61 NCDKSLHLTFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 85 NCDKSLHLTFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 180
Db 145 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 204

QY 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFPKAEFAVSKLVTDLT 240
Db 205 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFPKAEFAVSKLVTDLT 264
QY 241 VHTCCGGLLESCADDRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300
Db 265 VHTCCGGLLESCADDRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 324
QY 301 DPLSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYSVLLLRLLAKTYETTLTKC 360
Db 325 DPLSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYSVLLLRLLAKTYETTLTKC 384
QY 361 CAADPHCEYAKVDFEFPKPLVEEPQNLKONCELFEQLGEYKFNALLVRYTKVQPVYST 420
Db 385 CAADPHCEYAKVDFEFPKPLVEEPQNLKONCELFEQLGEYKFNALLVRYTKVQPVYST 444
QY 421 PTLVEVSRNLGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVEDTVYPKEFNAETFFHADICTLSEKEROIKKOTALVELVKHKPKAT 540
Db 505 LVNRRPCFSALEVEDTVYPKEFNAETFFHADICTLSEKEROIKKOTALVELVKHKPKAT 564
QY 541 KEQLKAVMDDFAAVFKECKCKADDKETCFABEGKLVAAASQAALGL 585
Db 565 KEQLKAVMDDFAAVFKECKCKADDKETCFABEGKLVAAASQAALGL 609

RESULT 33

US-08-897-956A-2
; Sequence 2, Application US/08897956A
; Patent No. 6423512
; GENERAL INFORMATION:
; APPLICANT: Mary Ellen Digan
; APPLICANT: Philip Lake
; APPLICANT: Hermann Gram
; TITLE OF INVENTION: Fusion Polypeptides
; FILE REFERENCE: 600-7244/CPA
; CURRENT APPLICATION NUMBER: US/08/997,956A
; CURRENT FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/022,689
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-08-897-956A-2

Query Match 82.7%; Score 484; DB 4; Length 609;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 584; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
Db 25 DAHSEVAHRFKDLGGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 84
QY 61 NCDKSLHLTFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 85 NCDKSLHLTFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 180
Db 145 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 204
QY 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFPKAEFAVSKLVTDLT 240
Db 205 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFPKAEFAVSKLVTDLT 264
QY 241 VHTCCGGLLESCADDRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300

Qy 181 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRPFAEYVSKLVTDLTK 240
Db 205 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRPFAEYVSKLVTDLTK 264
Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCECEKPLEKSHCIAEVENDEMPA 300
Db 265 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCECEKPLEKSHCIAEVENDEMPA 324
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMLFYEARHPDYSVLLLRLLAKTYETTTLEKC 360
Db 325 DLPSLAADFVESKDVCKNVAEAKDVLGMLFYEARHPDYSVLLLRLLAKTYETTTLEKC 384
Qy 361 CAAADPHECYAKVDFBPKLVEBPQNLIKQNCLEFQJGEYFQNALVRYTKVPQVST 420
Db 385 CAAADPHECYAKVDFBPKLVEBPQNLIKQNCLEFQJGEYFQNALVRYTKVPQVST 444
Qy 421 PTLVEVSRLNGVSKCKCKHPEAKMPCABDYL SVVLNQLCVLHKTVPVSRVTKCCTES 480
Db 445 PTLVEVSRLNGVSKCKCKHPEAKMPCABDYL SVVLNQLCVLHKTVPVSRVTKCCTES 504
Qy 481 LVNRRPCFSALEVDETYVPKFNFAETFTFHADICTLSEKEROIKKOTALVELVXHKPKAT 540
Db 505 LVNRRPCFSALEVDETYVPKFNFAETFTFHADICTLSEKEROIKKOTALVELVXHKPKAT 564
Qy 541 KEQLKAVMDDFAAVFEKCKKADDKETCFABEGKKLVAASQAALGL 585
Db 565 KEQLKAVMDDFAAVFEKCKKADDKETCFABEGKKLVAASQAALGL 609
RESULT 35
US-08-897-956A-3
; Sequence 3, Application US/08897956A
; Patent No. 6423512
; GENERAL INFORMATION:
; APPLICANT: Mary Ellen Digan
; APPLICANT: Philip Lake
; APPLICANT: Hermann Gram
; TITLE OF INVENTION: Fusion Polypeptides
; FILE REFERENCE: 800-7244/GPA
; CURRENT APPLICATION NUMBER: US/08/897,956A
; CURRENT FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/022,689
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 978
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion polypeptide
US-08-897-956A-3

Query Match 82.6%; Score 483; DB 4; Length 978;
Best Local Similarity 59.8%; Pred. No. 0;
Matches 583; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 DAHKEVAHRFKDLGLENFKALVLIAPAYLQOCPPEDHVKLVNEVTEFAKTCVADESAAE 60
Db 212 DAHKEVAHRFKDLGLENFKALVLIAPAYLQOCPPEDHVKLVNEVTEFAKTCVADESAAE 271
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREPV 120
Db 272 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREPV 331
Qy 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAAACLLP 180
Db 332 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAAACLLP 391
Qy 181 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRPFAEYVSKLVTDLTK 240
Db 392 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRPFAEYVSKLVTDLTK 451

Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCECEKPLEKSHCIAEVENDEMPA 300
Db 452 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCECEKPLEKSHCIAEVENDEMPA 511
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMLFYEARHPDYSVLLLRLLAKTYETTTLEKC 360
Db 512 DLPSLAADFVESKDVCKNVAEAKDVLGMLFYEARHPDYSVLLLRLLAKTYETTTLEKC 571
Qy 361 CAAADPHECYAKVDFBPKLVEBPQNLIKQNCLEFQJGEYFQNALVRYTKVPQVST 420
Db 572 CAAADPHECYAKVDFBPKLVEBPQNLIKQNCLEFQJGEYFQNALVRYTKVPQVST 631
Qy 421 PTLVEVSRLNGVSKCKCKHPEAKMPCABDYL SVVLNQLCVLHKTVPVSRVTKCCTES 480
Db 632 PTLVEVSRLNGVSKCKCKHPEAKMPCABDYL SVVLNQLCVLHKTVPVSRVTKCCTES 691
Qy 481 LVNRRPCFSALEVDETYVPKFNFAETFTFHADICTLSEKEROIKKOTALVELVXHKPKAT 540
Db 692 LVNRRPCFSALEVDETYVPKFNFAETFTFHADICTLSEKEROIKKOTALVELVXHKPKAT 751
Qy 541 KEQLKAVMDDFAAVFEKCKKADDKETCFABEGKKLVAASQAALG 584
Db 752 KEQLKAVMDDFAAVFEKCKKADDKETCFABEGKKLVAASQAALG 795
RESULT 36
US-10-153-064-102
; Sequence 102, Application US/101533064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 102
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (561)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-153-064-102

Query Match 81.7%; Score 478; DB 4; Length 668;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DAHKEVAHRFKDLGLENFKALVLIAPAYLQOCPPEDHVKLVNEVTEFAKTCVADESAAE 60
Db 83 DAHKEVAHRFKDLGLENFKALVLIAPAYLQOCPPEDHVKLVNEVTEFAKTCVADESAAE 142
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREPV 120
Db 143 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREPV 202
Qy 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAAACLLP 180
Db 203 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAAACLLP 262
Qy 181 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRPFAEYVSKLVTDLTK 240
Db 263 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSQRPFAEYVSKLVTDLTK 322
Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCECEKPLEKSHCIAEVENDEMPA 300
Db 323 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCECEKPLEKSHCIAEVENDEMPA 382

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFYARRHPDYSVVLRLAKTYETTLK 360
DB 383 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFYARRHPDYSVVLRLAKTYETTLK 442
QY 361 CAADAPHECYAKVDFEKPFLVEBPQNLKONCELFEQGEYKFNQALLVRYTKVPQVST 420
DB 443 CAADAPHECYAKVDFEKPFLVEBPQNLKONCELFEQGEYKFNQALLVRYTKVPQVST 502
QY 421 PTLVEVSRNLGKVGSKCKGPKAPKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCT 478
DB 503 PTLVEVSRNLGKVGSKCKGPKAPKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCT 560

RESULT 37

US-10-153-064-101
; Sequence 101, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PFS56
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 101
; LENGTH: 692
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (585)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-153-064-101

Query Match 81.7%; Score 478; DB 4; Length 692;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADES 60
DB 107 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADES 166
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 167 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 226
QY 121 DVMCTAFHDNEETFLKKLYEIAARRHPYFPAPELLFFAKRYKAAFTCCQAAADKAACLLP 180
DB 227 DVMCTAFHDNEETFLKKLYEIAARRHPYFPAPELLFFAKRYKAAFTCCQAAADKAACLLP 286
QY 181 KLDELDEGKASSAKQRLKCSLOKGERAFKAWAVARLSQRPFAEFVSKLVTDLTK 240
DB 287 KLDELDEGKASSAKQRLKCSLOKGERAFKAWAVARLSQRPFAEFVSKLVTDLTK 346
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPFLKSHCIAEVENDEMPA 300
DB 347 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPFLKSHCIAEVENDEMPA 406
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFYARRHPDYSVVLRLAKTYETTLK 360
DB 407 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFYARRHPDYSVVLRLAKTYETTLK 466
QY 361 CAADAPHECYAKVDFEKPFLVEBPQNLKONCELFEQGEYKFNQALLVRYTKVPQVST 420
DB 467 CAADAPHECYAKVDFEKPFLVEBPQNLKONCELFEQGEYKFNQALLVRYTKVPQVST 526
QY 421 PTLVEVSRNLGKVGSKCKGPKAPKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCT 478
DB 527 PTLVEVSRNLGKVGSKCKGPKAPKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCT 584

RESULT 38
US-08-448-196A-3
; Sequence 3, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
; TITLE OF INVENTION: RELATED PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NASA
; STREET: MARSHALL SPACE FLIGHT CENTER
; CITY: HUNTSVILLE
; STATE: ALABAMA
; COUNTRY: USA
; ZIP: 35812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,196A
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: BROAD JR., ROBERT L.
; REGISTRATION NUMBER: 18,757
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 205-544-0021
; TELEFAX: 205-544-0258
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-448-196A-3

Query Match 65.5%; Score 383; DB 1; Length 585;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 583; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADES 60
DB 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADES 60
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKLYEIAARRHPYFPAPELLFFAKRYKAAFTCCQAAADKAACLLP 180
DB 121 DVMCTAFHDNEETFLKKLYEIAARRHPYFPAPELLFFAKRYKAAFTCCQAAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKCSLOKGERAFKAWAVARLSQRPFAEFVSKLVTDLTK 240
DB 181 KLDELDEGKASSAKQRLKCSLOKGERAFKAWAVARLSQRPFAEFVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPFLKSHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPFLKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFYARRHPDYSVVLRLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFYARRHPDYSVVLRLAKTYETTLK 360

QY 361 CAADPHCYAKVDFBFLVPEEPONLQKONCELFQELGEYKFNQALLVRYTKVPQVST 420
DB 361 CAADPHCYAKVDFBFLVPEEPONLQKONCELFQELGEYKFNQALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVHKEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVHKEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETVVPKFNFAETFTFHADICTLSEKERQIKKQATLVKHKPKAT 540
DB 481 LVNRRPCFSALEVDETVVPKFNFAETFTFHADICTLSEKERQIKKQATLVKHKPKAT 540
QY 541 KEQLKAVMDPFAAFVEKCKKADDDKTCFAEKGKLVAAASQAALGL 585
DB 541 KEQLKAVMDPFAAFVEKCKKADDDKTCFAEKGKLVAAASQAALGL 585
RESULT 39
US-08-984-176-1
; Sequence 1, Application US/08984176
; Patent No. 5948609
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C
; APPLICANT: HO, JOSEPH X
; APPLICANT: RUKER, FLORIAN
; TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT
; TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER
; FILE REFERENCE: 08/984,176
; CURRENT APPLICATION NUMBER: US/08/984,176
; CURRENT FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-984-176-1
Query Match 65.5%; Score 383; DB 2; Length 585;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 583; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 DAHKEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCAVDESAB 60
DB 1 DAHKEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCAVDESAB 60
QY 61 NCDKSLHTLFGDKLCTVATLETYGEWADCCAKOEPERNECFLOHKKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLETYGEWADCCAKOEPERNECFLOHKKDDNPNLRLVRPEV 120
QY 121 DVWCTAFHDNEETFLKKYLYEIAIRRHYPFYAPPELLFFAKRYKAFTCCQAADKAACLLP 180
DB 121 DVWCTAFHDNEETFLKKYLYEIAIRRHYPFYAPPELLFFAKRYKAFTCCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKORLKASLOKFGERAPKANAVARLSORFPKAEFAVSKLVTDLT 240
DB 181 KLDELDEGKASSAKORLKASLOKFGERAPKANAVARLSORFPKAEFAVSKLVTDLT 240
QY 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLKCECKPPLLEKSHCIAEVENDEMPA 300
DB 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLKCECKPPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADPVESKDYCKNYAEAKOVFLGMFLYEVARRHDPYSVLLRLAKTVEITTEK 360
DB 301 DLPSLAADPVESKDYCKNYAEAKOVFLGMFLYEVARRHDPYSVLLRLAKTVEITTEK 360
QY 361 CAADPHCYAKVDFBFLVPEEPONLQKONCELFQELGEYKFNQALLVRYTKVPQVST 420
DB 361 CAADPHCYAKVDFBFLVPEEPONLQKONCELFQELGEYKFNQALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVHKEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVHKEKTPVSDRVTKCCTES 480

QY 481 LVNRRPCFSALEVDETVVPKFNFAETFTFHADICTLSEKERQIKKQATLVKHKPKAT 540
DB 481 LVNRRPCFSALEVDETVVPKFNFAETFTFHADICTLSEKERQIKKQATLVKHKPKAT 540
QY 541 KEQLKAVMDPFAAFVEKCKKADDDKTCFAEKGKLVAAASQAALGL 585
DB 541 KEQLKAVMDPFAAFVEKCKKADDDKTCFAEKGKLVAAASQAALGL 585
RESULT 40
US-09-846-329A-1
; Sequence 1, Application US/09846329A
; Patent No. 6630786
; GENERAL INFORMATION:
; APPLICANT: Jackowski, George
; TITLE OF INVENTION: Biopolymer Marker Indicative of Disease State Having A Molecular
; TITLE OF INVENTION: OF 2937 Daltons
; FILE REFERENCE: 2132.052
; CURRENT APPLICATION NUMBER: US/09/846,329A
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 1
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-846-329A-1
Query Match 4.6%; Score 27; DB 4; Length 28;
Best Local Similarity 100.0%; Pred. No. 9.4e-19;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKEVAHRFKDLGEENFKALVLIAP 27
DB 2 DAHKEVAHRFKDLGEENFKALVLIAP 28
RESULT 41
US-08-448-196A-5
; Sequence 5, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
; TITLE OF INVENTION: RELATED PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NASA
; STREET: MARSHALL SPACE FLIGHT CENTER
; CITY: HUNTSVILLE
; STATE: ALABAMA
; COUNTRY: USA
; ZIP: 35812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,196A
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: BROAD JR., ROBERT L.
; REGISTRATION NUMBER: 18,757
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 205-544-0021
; TELEFAX: 205-544-0258
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:

```
; LENGTH: 583 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-448-196A-5

Query Match          4.4%; Score 27; DB 1; Length 583;
Best Local Similarity 100.0%; Pred. No. 1.3e-17;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 ADESAENCDSLHTLFGDKLCTVATLR 81
DB 55 ADESAENCDSLHTLFGDKLCTVATLR 81

RESULT 42
US-08-448-196A-7
; Sequence 7, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
; RELATED PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NASA
; STREET: MARSHALL SPACE FLIGHT CENTER
; CITY: HUNTSVILLE
; STATE: ALABAMA
; COUNTRY: USA
; ZIP: 35812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,196A
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: BROAD JR., ROBERT L.
; REGISTRATION NUMBER: 18,757
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 205-544-0021
; TELEFAX: 205-544-0258
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 584 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-448-196A-1

Query Match          4.4%; Score 26; DB 1; Length 584;
Best Local Similarity 100.0%; Pred. No. 1.3e-16;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 ADCCAKQEPERNECFLOKDDNPILP 113
DB 88 ADCCAKQEPERNECFLOKDDNPILP 113

RESULT 43
US-08-448-196A-1
; Sequence 1, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
; RELATED PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NASA
; STREET: MARSHALL SPACE FLIGHT CENTER
; CITY: HUNTSVILLE
; STATE: ALABAMA
; COUNTRY: USA
; ZIP: 35812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,196A
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: BROAD JR., ROBERT L.
; REGISTRATION NUMBER: 18,757
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 205-544-0021
; TELEFAX: 205-544-0258
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 584 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-448-196A-7

Query Match          3.8%; Score 22; DB 1; Length 109;
Best Local Similarity 100.0%; Pred. No. 2.2e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 ECCHGDLLECCADRDADLAKYIC 265
DB 54 ECCHGDLLECCADRDADLAKYIC 75

RESULT 44
US-08-134-638-1
; Sequence 1, Application US/08134638
; Patent No. 5473050
; GENERAL INFORMATION:
; APPLICANT: Strand, Frederick T
; TITLE OF INVENTION: Denatured Bovine Serum Albumin Milk
; PRODUCTS AND METHOD THEREFOR
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Frederick T. Strand
; STREET: P.O. Box 64321
; CITY: Phoenix
; STATE: Arizona
; COUNTRY: USA
; ZIP: 85082-4321
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 1.2 Mb storage
; COMPUTER: IBM PC
; OPERATING SYSTEM: MS-DOS 5.0
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/134,638
```

Query Match 3.8%; Score 22; DB 1; Length 583;
Best Local Similarity 100.0%; Pred. No. 9.8e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 ECCHGDLLECCADRADLAKYIC 265
Db 243 ECCHGDLLECCADRADLAKYIC 264

RESULT 46
US-08-448-196A-6
; Sequence 6, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
; TITLE OF INVENTION: RELATED PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NASA
; STREET: MARSHALL SPACE FLIGHT CENTER
; CITY: HUNTSVILLE
; STATE: ALABAMA
; COUNTRY: USA
; ZIP: 35812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,196A
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: BROAD JR., ROBERT L.
; REGISTRATION NUMBER: 18,757
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 205-544-0021
; TELEFAX: 205-544-0258
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 583 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-448-196A-6

Query Match 3.8%; Score 22; DB 1; Length 583;
Best Local Similarity 100.0%; Pred. No. 9.8e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 ECCHGDLLECCADRADLAKYIC 265
Db 243 ECCHGDLLECCADRADLAKYIC 264

RESULT 47
US-08-448-196A-2
; Sequence 2, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
; TITLE OF INVENTION: RELATED PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:

FILING DATE: 10/12/93
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: N/A
FILING DATE: N/A
ATTORNEY/AGENT INFORMATION:
NAME: Weiss, Harry M
REGISTRATION NUMBER: 19,497
REFERENCE/DOCKET NUMBER: 1795P1423
TELECOMMUNICATION INFORMATION:
TELEPHONE: (602) 994-8988
TELEFAX: (602) 947-2663
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 582
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-134-638-1

Query Match 3.8%; Score 22; DB 1; Length 582;
Best Local Similarity 100.0%; Pred. No. 9.8e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 ECCHGDLLECCADRADLAKYIC 265
Db 242 ECCHGDLLECCADRADLAKYIC 263

RESULT 45
US-08-448-196A-4
; Sequence 4, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
; TITLE OF INVENTION: RELATED PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NASA
; STREET: MARSHALL SPACE FLIGHT CENTER
; CITY: HUNTSVILLE
; STATE: ALABAMA
; COUNTRY: USA
; ZIP: 35812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,196A
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: BROAD JR., ROBERT L.
; REGISTRATION NUMBER: 18,757
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 205-544-0021
; TELEFAX: 205-544-0258
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 583 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-448-196A-4

ADDRESSEE: NASA
STREET: MARSHALL SPACE FLIGHT CENTER
CITY: HUNTSVILLE
STATE: ALABAMA
COUNTRY: USA
ZIP: 35812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,196A
FILING DATE: 23-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: BROAD JR., ROBERT L.
REGISTRATION NUMBER: 18,757
REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 205-544-0021
TELEFAX: 205-544-0258
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-448-196A-2

Query Match 3.2%; Score 19; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 1.8e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 473 VTKCCTESLVNRRPCFSAL 491
DB 95 VTKCCTESLVNRRPCFSAL 113

RESULT 48
US-08-470-187-1
Sequence 1, Application US/08470187
Patent No. 5532152
GENERAL INFORMATION:
APPLICANT: Cousens, Lawrence S.
APPLICANT: Eberhardt, Christine E.
APPLICANT: Gray, Patrick W.
APPLICANT: Tjoelker, Larry W.
APPLICANT: Wilder, Cheryl L.
TITLE OF INVENTION: Platelet-Activating Factor Acetyl
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gertein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,187
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: No. 5532152and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31672
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3658
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-470-187-1
Query Match 2.9%; Score 17; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 FKDLGEENFKALVLIAP 27
DB 1 FKDLGEENFKALVLIAP 17

RESULT 49
US-08-318-905-1
Sequence 1, Application US/08318905
Patent No. 5641669
GENERAL INFORMATION:
APPLICANT: Cousens, Lawrence S.
APPLICANT: Eberhardt, Christine D.
APPLICANT: Gray, Patrick W.
APPLICANT: Le Trong, Hai
APPLICANT: Tjoelker, Larry W.
APPLICANT: Wilder, Cheryl L.
TITLE OF INVENTION: Platelet-Activating Factor Acetyl
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gertein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,905
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/133,803
FILING DATE: 6-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: No. 5641669and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 32205
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3658
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide

US-08-318-905-1

Query Match 2.9%; Score 17; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 FKDLGEENFKALVLIAP 27
Db 1 FKDLGEENFKALVLIAP 17

RESULT 50

US-08-483-232-1
Sequence 1, Application US/08483232
Patent No. 5656431
GENERAL INFORMATION:
APPLICANT: Cousens, Lawrence S.
APPLICANT: Eberhardt, Christine D.
APPLICANT: Gray, Patrick W.
APPLICANT: Le Trong, Hai
APPLICANT: Tjoelker, Larry W.
APPLICANT: Wilder, Cheryl L.
TITLE OF INVENTION: Platelet-Activating Factor
TITLE OF INVENTION: Acetylhydrolase
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/483,232
FILING DATE: 06-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,905
FILING DATE: 06-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: No. 5656431and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/32689
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3658
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide

Query Match 2.9%; Score 17; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 FKDLGEENFKALVLIAP 27
Db 1 FKDLGEENFKALVLIAP 17

RESULT 51

US-08-483-140-1

Sequence 1, Application US/08483140
Patent No. 5696403
GENERAL INFORMATION:
APPLICANT: ICOS Corporation
TITLE OF INVENTION: Platelet-Activating Factor Acetyl
TITLE OF INVENTION: Hydrolase
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/483,140
FILING DATE: 06-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,905
FILING DATE: 06-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: No. 5696403and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 32781
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3658
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide

Query Match 2.9%; Score 17; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 FKDLGEENFKALVLIAP 27
Db 1 FKDLGEENFKALVLIAP 17

RESULT 52

US-08-485-938A-1
Sequence 1, Application US/08485938A
Patent No. 5847088
GENERAL INFORMATION:
APPLICANT: Cousens, Lawrence S.
APPLICANT: Eberhardt, Christine D.
APPLICANT: Gray, Patrick W.
APPLICANT: Le Trong, Hai
APPLICANT: Tjoelker, Larry W.
APPLICANT: Wilder, Cheryl L.
TITLE OF INVENTION: Platelet-Activating Factor
TITLE OF INVENTION: Acetylhydrolase
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago

STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,938A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA: US 08/318,905
FILING DATE: 06-OCT-1994
APPLICATION NUMBER: US 08/133,803
FILING DATE: 06-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: No. 5847088and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/32792
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3658
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-485-938A-1

Query Match 2.9%; Score 17; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
QY 11 FKDLGEENFKALVLIAP 27
Db 1 FKDLGEENFKALVLIAP 17
RESULT 53
US-08-910-041-1
Sequence 1, Application US/08910041
Patent No. 5977308
GENERAL INFORMATION:
APPLICANT: Cousens, Lawrence S.
APPLICANT: Eberhardt, Christine D.
APPLICANT: Gray, Patrick W.
APPLICANT: Le Trong, Hai
APPLICANT: Tjoelker, Larry W.
APPLICANT: Wilder, Cheryl L.
TITLE OF INVENTION: Platelet-Activating Factor
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910,041
FILING DATE:
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/483,232
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,905
FILING DATE: 06-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/133,803
FILING DATE: 06-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/34026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3658
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-910-041-1
Query Match 2.9%; Score 17; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
QY 11 FKDLGEENFKALVLIAP 27
Db 1 FKDLGEENFKALVLIAP 17
RESULT 54
US-09-328-474-1
Sequence 1, Application US/09328474
Patent No. 6045794
GENERAL INFORMATION:
APPLICANT: Cousens, Lawrence S.
APPLICANT: Eberhardt, Christine D.
APPLICANT: Gray, Patrick W.
APPLICANT: Le Trong, Hai
APPLICANT: Tjoelker, Larry W.
APPLICANT: Wilder, Cheryl L.
TITLE OF INVENTION: Platelet-Activating Factor
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/328,474
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/483,232
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,905
FILING DATE: 06-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/133,803
FILING DATE: 06-OCT-1993

ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/34026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3658
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-328-474-1

Query Match 2.9%; Score 17; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 FKDLGEENFKALVLI AF 27
Db 1 FKDLGEENFKALVLI AF 17

RESULT 55

US-09-100-546-1
Sequence 1, Application US/09100546
Patent No. 609836
GENERAL INFORMATION:
APPLICANT: Cousens, Lawrence S.
APPLICANT: Eberhardt, Christine D.
APPLICANT: Gray, Patrick W.
APPLICANT: Le Trong, Hai
APPLICANT: Tjoelker, Larry W.
TITLE OF INVENTION: Platelet-Activating Factor
TITLE OF INVENTION: Acetylhydrolase
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,546
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/010,715
FILING DATE:
APPLICATION NUMBER: US 08/318,905
FILING DATE: 06-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/133,803
FILING DATE: 06-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: No. 609836and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/32793
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3658
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-328-474-1

Query Match 2.9%; Score 17; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 FKDLGEENFKALVLI AF 27
Db 1 FKDLGEENFKALVLI AF 17

LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-546-1

Query Match 2.9%; Score 17; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 FKDLGEENFKALVLI AF 27
Db 1 FKDLGEENFKALVLI AF 17

RESULT 56

US-09-010-715-1
Sequence 1, Application US/09010715
Patent No. 6146625
GENERAL INFORMATION:
APPLICANT: Cousens, Lawrence S.
APPLICANT: Eberhardt, Christine D.
APPLICANT: Gray, Patrick W.
APPLICANT: Le Trong, Hai
APPLICANT: Tjoelker, Larry W.
APPLICANT: Wilder, Cheryl L.
TITLE OF INVENTION: Platelet-Activating Factor
TITLE OF INVENTION: Acetylhydrolase
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/010,715
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,905
FILING DATE: 06-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/133,803
FILING DATE: 06-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: No. 6146625and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/32793
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3658
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-010-715-1

Query Match 2.9%; Score 17; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 FKDLGEENFKALVLI AF 27
Db 1 FKDLGEENFKALVLI AF 17

Db 1 FKDLGEENFKALVLIIF 17

RESULT 57

US-09-577-758-1
; Sequence 1, Application US/09577758
; Patent No. 6203750
; GENERAL INFORMATION:
; APPLICANT: Cousens, Lawrence S.
; APPLICANT: Eberhardt, Christine D.
; APPLICANT: Gray, Patrick W.
; APPLICANT: Le Trong, Hai
; APPLICANT: Tisler, Larry W.
; APPLICANT: Wilder, Cheryl L.
; TITLE OF INVENTION: Platelet-Activating Factor
; TITLE OF INVENTION: Acetylhydrolase
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/577,758
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/010,715
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/133,803
; FILING DATE: 06-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6203790and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/32793
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3658
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-577-758-1

Query Match 2.9%; Score 17; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 FKDLGEENFKALVLIIF 27

Db 1 FKDLGEENFKALVLIIF 17

RESULT 58

US-08-702-572-16
; Sequence 16, Application US/08702572
; Patent No. 5965386
; GENERAL INFORMATION:
; APPLICANT: Kerry-Williams, Sean M
; APPLICANT: Gilbert, Sarah C
; TITLE OF INVENTION: Yeast Strains and Modified Albumins
; NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:
ADDRESSEE: Centeon L.L.C.
STREET: 1020 First Avenue
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-1310
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,572
FILING DATE: 11-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 95/23857
FILING DATE: 1-MAR-1995
APPLICATION NUMBER: GB 9404270.2
FILING DATE: 5-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Naomi Biswas
REGISTRATION NUMBER: 38,384
REFERENCE/DOCKET NUMBER: CE0114 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610/878/4294
TELEFAX: 610/878/4221
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
US-08-702-572-16

Query Match 2.6%; Score 15; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 403 FQNALVRYTKVPQ 417

Db 1 FQNALVRYTKVPQ 15

RESULT 59

US-10-053-485-24
; Sequence 24, Application US/10053485
; Patent No. 8578896
; GENERAL INFORMATION:
; APPLICANT: Figeys, Daniel
; APPLICANT: Abersold, Ruedi
; TITLE OF INVENTION: ELECTROSMOTIC FLUIDIC DEVICE AND RELATED METHODS
; FILE REFERENCE: UWO1118617
; CURRENT APPLICATION NUMBER: US/10/053,485
; CURRENT FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: US 09/209,880
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: US 60/069,398
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-053-485-24

Query Match

2.6%; Score 15; DB 4; Length 16;

Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 414 KVPQSTPTLVEVSR 428
Db 2 KVPQSTPTLVEVSR 16

RESULT 60

US-08-787-547-30
; Sequence 30, Application US/08787547
; Patent No. 5783567
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; APPLICANT: Curley, Joanne M.
; APPLICANT: Langer, Robert S.
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY
; TITLE OF INVENTION: OF NUCLEIC ACID
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/787,547
; FILING DATE: 22-JAN-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 08191/003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-787-547-30

Query Match 2.2%; Score 13; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 420 TPTLVEVSRNLGK 432
Db 1 TPTLVEVSRNLGK 13

RESULT 61

US-08-803-364-7
; Sequence 7, Application US/08803364
; Patent No. 5864014
; GENERAL INFORMATION:
; APPLICANT: PASANO, Alessio
; TITLE OF INVENTION: ZONULA OCCLUDENS TOXIN RECEPTOR
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS

STREET: 2100 Pennsylvania Avenue, N.W., Suite 800
CITY: Washington, D.C.
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/803,364
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/_____
; FILING DATE: 20 FEB 1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kit, Gordon
; REGISTRATION NUMBER: 30,764
; REFERENCE/DOCKET NUMBER: A-6888
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
US-08-803-364-7

Query Match 2.2%; Score 13; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AHKSEVAHRFKDL 14
Db 1 AHKSEVAHRFKDL 13

RESULT 62

US-09-024-198-13
; Sequence 13, Application US/09024198
; Patent No. 5912323
; GENERAL INFORMATION:
; APPLICANT: PASANO, Alessio
; TITLE OF INVENTION: ZONULA OCCLUDENS TOXIN RECEPTOR
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W., Suite 800
; CITY: Washington, D.C.
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/024,198
; FILING DATE: 17 FEB 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/803,364
; FILING DATE: 20 FEB 1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kit, Gordon

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AHKSEVAHRFKDL 14
| | | | | | | | | | | | | | |
DB 1 AHKSEVAHRFKDL 13

RESULT 64
US-08-153-799-10
; Sequence 10, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/153,799
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847975
FILING DATE: 06-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8909916.2
FILING DATE: 29-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB90/00650
FILING DATE: 26-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/775952
FILING DATE: 29-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Swope, R Hain
REGISTRATION NUMBER: 24864
REFERENCE/DOCKET NUMBER: 92H832
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 665 2400
TELEFAX: (908) 771 6159
TELEX: 219484
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-153-799-10

Query Match 2.1%; Score 12; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00014;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 DELRDEGKASSA 194
| | | | | | | | | | | | | | |
DB 1 DELRDEGKASSA 12

RESULT 65
US-10-053-485-26
; Sequence 26, Application US/10053485
; Patent No. 6576896

; REGISTRATION NUMBER: 30,764
; REFERENCE/DOCKET NUMBER: A-6988
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
; HYPOTHETICAL: NO
US-09-024-198-13

Query Match 2.2%; Score 13; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AHKSEVAHRFKDL 14
| | | | | | | | | | | | | | |
DB 1 AHKSEVAHRFKDL 13

RESULT 63
US-09-186-409-13
; Sequence 13, Application US/09186409
; Patent No. 5948629
; GENERAL INFORMATION:
; APPLICANT: FASANO, Alessio
; TITLE OF INVENTION: ZONULA OCCLUDENS TOXIN RECEPTOR
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W., Suite 800
; CITY: Washington, D.C.
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/186,409
FILING DATE: 5 NOV 1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/024,198
FILING DATE: 17 FEB 1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/803,364
FILING DATE: 20 FEB 1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: KIT, Gorgon
REGISTRATION NUMBER: 30,764
REFERENCE/DOCKET NUMBER: A-7306
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Peptide
HYPOTHETICAL: NO
US-09-186-409-13

Query Match 2.2%; Score 13; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;

GENERAL INFORMATION:
APPLICANT: Figeys, Daniel
TITLE OF INVENTION: ELECTROSMOTIC FLUIDIC DEVICE AND RELATED METHODS
FILE REFERENCE: UW01118617
CURRENT APPLICATION NUMBER: US/10/053,485
CURRENT FILING DATE: 2002-05-28
PRIOR APPLICATION NUMBER: US 09/209,880
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: US 60/069,398
PRIOR FILING DATE: 1997-12-12
NUMBER OF SEQ ID NOS: 66
SOFTWARE: PatentIn version 3.0
SEQ ID NO 26
LENGTH: 17
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-10-053-485-26

Query Match 2.1%; Score 12; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 RRHPYFAPELL 155
DB 1 RRHPYFAPELL 12

RESULT 66
US-08-378-859-4
Sequence 4, Application US/08378859
Patent No. 5728553
GENERAL INFORMATION:
APPLICANT: Goodey, Andrew R
APPLICANT: Sleep, Darrell
APPLICANT: van Urk, Hendrik
APPLICANT: Berezenko, Stephen
APPLICANT: Woodrow, John R
APPLICANT: Johnson, Richard A
APPLICANT: Wood, Patricia C
APPLICANT: Burton, Steven J
APPLICANT: Quirk, Alan V
TITLE OF INVENTION: High Purity Albumin
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: The BOC Group, Inc
STREET: 100 Mountain Avenue
CITY: Murray Hill, New Providence
STATE: New Jersey
COUNTRY: United States of America
ZIP: 07974
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/378,859
FILING DATE: 26-JAN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Swope, R Hain
REGISTRATION NUMBER: 24,864
REFERENCE/DOCKET NUMBER: 92H8501
TELEPHONE: (908) 771 6159
TELEFAX: (908) 771 6292
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Peptide fragment of human serum albumin
US-08-378-859-4

Query Match 1.9%; Score 11; DB 1; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 575 LVAAQAALGL 585
DB 1 LVAAQAALGL 11

RESULT 67
US-08-970-648-4
Sequence 4, Application US/08970648
Patent No. 6034221
GENERAL INFORMATION:
APPLICANT: Goodey, Andrew R
APPLICANT: Sleep, Darrell
APPLICANT: van Urk, Hendrik
APPLICANT: Berezenko, Stephen
APPLICANT: Woodrow, John R
APPLICANT: Johnson, Richard A
APPLICANT: Wood, Patricia C
APPLICANT: Burton, Steven J
APPLICANT: Quirk, Alan V
TITLE OF INVENTION: High Purity Albumin
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: The BOC Group, Inc
STREET: 100 Mountain Avenue
CITY: Murray Hill, New Providence
STATE: New Jersey
COUNTRY: United States of America
ZIP: 07974
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/970,648
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/378,859
FILING DATE: 26-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Swope, R Hain
REGISTRATION NUMBER: 24,864
REFERENCE/DOCKET NUMBER: 92H8501
TELEPHONE: (908) 771 6159
TELEFAX: (908) 771 6292
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Peptide fragment of human serum albumin
US-08-970-648-4

Qy 145 RHPYFYAPELL 155

APPLICANT: Berezhenko, Stephen
APPLICANT: Woodrow, John R
APPLICANT: Johnson, Richard A
APPLICANT: Wood, Patricia C
APPLICANT: Burton, Steven J
APPLICANT: Quirk, Alan V
TITLE OF INVENTION: High Purity Albumin
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: The BOC Group, Inc
STREET: 100 Mountain Avenue
CITY: Murray Hill, New Providence
STATE: New Jersey
COUNTRY: United States of America
ZIP: 07974
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/970,648
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/378,859
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Swope, R Hain
REGISTRATION NUMBER: 24,864
REFERENCE/DOCKET NUMBER: 92H8501
TELEPHONE: (908) 771 6292
TELEFAX: (908) 771 6159
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Peptide fragment of human serum albumin
US-08-970-648-1

Query Match 1.7%; Score 10; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 575 LVAASQAALG 584
Db 1 LVAASQAALG 10

RESULT 72
US-08-952-558-2
Sequence 2, Application US/08952558
Patent No. 6638740
GENERAL INFORMATION:
APPLICANT: Goodey, Andrew R.
APPLICANT: Sleep, Darrell
APPLICANT: Berezhenko, Stephen
APPLICANT: Woodrow, John R.
APPLICANT: Johnson, Richard A.
TITLE OF INVENTION: PROCESS OF HIGH PURITY ALBUMIN PRODUCTION
FILE REFERENCE: CE0111D US
CURRENT APPLICATION NUMBER: US/08/952,558
CURRENT FILING DATE: 1999-02-24
PRIOR APPLICATION NUMBER: PCT/GB96/00449
PRIOR FILING DATE: 1996-02-29

PRIOR APPLICATION NUMBER: 378,859
PRIOR FILING DATE: 1995-05-25
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 10
TYPE: PRT
ORGANISM: Homo sapiens
US-08-952-558-2
Query Match 1.7%; Score 10; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 575 LVAASQAALG 584
Db 1 LVAASQAALG 10
RESULT 73
US-08-469-856-9
Sequence 9, Application US/08469856
Patent No. 5650307
GENERAL INFORMATION:
APPLICANT: Sijmons, Peter C.
APPLICANT: Hoekema, Andreas
APPLICANT: Dekker, Bernardus M.M.
APPLICANT: Schrammeijer, Barbara
APPLICANT: Verwoerd, Tewis C.
APPLICANT: Vandeneizen, Petrus J.M.
TITLE OF INVENTION: PRODUCTION OF HETEROLOGOUS PROTEINS IN
TITLE OF INVENTION: PLANTS AND PLANT CELLS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Avenue, NW, Ste. 5500
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,856
FILING DATE: 06-JUN-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 6192-0025.01
TELEPHONE: (202)887-1500
TELEFAX: (202)887-0763
TELEX: 90-4030 MRSNFOERSWSH
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-469-856-9

Query Match 1.5%; Score 9; DB 1; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DAHKSEVAH 9
Db 3 DAHKSEVAH 11

RESULT 74

US-08-378-859-3
Sequence 3, Application US/08378859
Patent No. 5728553

GENERAL INFORMATION:

APPLICANT: Goodey, Andrew R
APPLICANT: Sleep, Darrell
APPLICANT: van Urk, Hendrik
APPLICANT: Berezhenko, Stephen
APPLICANT: Woodrow, John R
APPLICANT: Johnson, Richard A
APPLICANT: Wood, Patricia C
APPLICANT: Burton, Steven J
APPLICANT: Quirk, Alan V
TITLE OF INVENTION: High Purity Albumin
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: The BOC Group, Inc
STREET: 100 Mountain Avenue
CITY: Murray Hill, New Providence
STATE: New Jersey
COUNTRY: United States of America
ZIP: 07974

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/378,859
FILING DATE: 26-JAN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Swope, R Hain
REGISTRATION NUMBER: 24,864
REFERENCE/DOCKET NUMBER: 92H8501
TELEPHONE: (908) 771 6292
TELEFAX: (908) 771 6159
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE: Peptide fragment of human serum albumin

US-08-378-859-3

Query Match 1.4%; Score 8; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 DLGEENFK 20

Db 1 DLGEENFK 8

RESULT 75

US-08-370-648-3
Sequence 3, Application US/08970648
Patent No. 6034221

GENERAL INFORMATION:

APPLICANT: Goodey, Andrew R
APPLICANT: Sleep, Darrell
APPLICANT: van Urk, Hendrik
APPLICANT: Berezhenko, Stephen
APPLICANT: Woodrow, John R

US-08-378-859-3

Query Match 1.4%; Score 8; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 DLGEENFK 20

Db 1 DLGEENFK 8

APPLICANT: Johnson, Richard A
APPLICANT: Wood, Patricia C
APPLICANT: Burton, Steven J
APPLICANT: Quirk, Alan V
TITLE OF INVENTION: High Purity Albumin
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: The BOC Group, Inc
STREET: 100 Mountain Avenue
CITY: Murray Hill, New Providence
STATE: New Jersey
COUNTRY: United States of America
ZIP: 07974
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/970,648
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/378,859
FILING DATE: 26-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Swope, R Hain
REGISTRATION NUMBER: 24,864
REFERENCE/DOCKET NUMBER: 92H8501
TELEPHONE: (908) 771 6292
TELEFAX: (908) 771 6159
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE: Peptide fragment of human serum albumin

US-08-970-648-3

Query Match 1.4%; Score 8; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 DLGEENFK 20

Db 1 DLGEENFK 8

Search completed: April 19, 2004, 16:08:26
Job time : 26 secs

CC therapeutic protein and albumin (e.g. human serum albumin (HSA or HA)).
CC The albumin fusion proteins are useful for treating, preventing, or
CC ameliorating various disorders. Such disorders include immune disorders,
CC autoimmune disorders, blood-related disorders, hyperproliferative
CC disorders, renal disorders, cardiovascular disorders, respiratory
CC disorders, neurological disorders, endocrine disorders, reproductive
CC system disorders, gastrointestinal disorders, infectious disease, and
CC wound healing. Therapeutic proteins can be stabilised to extend shelf
CC life and/or retain the protein's activity for extended periods of time in
CC solution, in vivo or in vitro by genetically or chemically fusing the
CC protein to albumin or its fragment or variant. In addition the use of
CC albumin fusion proteins reduces the need to formulate protein solutions
CC with large excesses of carrier proteins to prevent loss of therapeutic
CC protein due to factors such as binding to the container. The extension of
CC shelf life was tested by measuring biological activity (Nb2 cell
CC proliferation) of human albumin-human growth hormone (HA-hGH) fusion
CC protein remaining after incubation in cell culture media for up to 3
CC weeks. At week 3 there was still approximately 95% cell proliferation
CC compared to no activity of unfused hGH. The present sequence represents
CC the mature form of HSA which can be used to produce the albumin fusion
CC proteins of the invention
XX
XX
SQ Sequence 585 AA;

Query Match 100.0%; Score 585; DB 5; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFKDLGSENFKALVIAFAQYLOQCFFEDHVKLVNEVTEFAKTCVADSEAE 60
DB 1 DAHKSVAHRFKDLGSENFKALVIAFAQYLOQCFFEDHVKLVNEVTEFAKTCVADSEAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLPLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLPLVRPEV 120

QY 121 DVNCTAFHNEETFLKKLYEIAARRHPFYAPPELLFFAKRYKAAPTECCQADKAACLLP 180
DB 121 DVNCTAFHNEETFLKKLYEIAARRHPFYAPPELLFFAKRYKAAPTECCQADKAACLLP 180

QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRFPPKAEFAEYVKLVTDLT 240
DB 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRFPPKAEFAEYVKLVTDLT 240

QY 241 VHTCECHGDLLECADRADLAKYICENQDISISKLECEKPLEKSHCIAEVENDEMPA 300
DB 241 VHTCECHGDLLECADRADLAKYICENQDISISKLECEKPLEKSHCIAEVENDEMPA 300

QY 301 DLPSLAADFVSKDVCNVAEAKDVLGMFLVEYARRHPDYSVLLLLRLAKTYETTLK 360
DB 301 DLPSLAADFVSKDVCNVAEAKDVLGMFLVEYARRHPDYSVLLLLRLAKTYETTLK 360

QY 361 CAADPHECYAKVDFEKPVEEPQNLKONCELFELQGEYKFNALLVRYTKVPQYST 420
DB 361 CAADPHECYAKVDFEKPVEEPQNLKONCELFELQGEYKFNALLVRYTKVPQYST 420

QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

QY 481 LVNRRPCFSALEVDVTVKPEFNAETTFHADICTLSEKERQIKKOTALVELVKHKPKAT 540
DB 481 LVNRRPCFSALEVDVTVKPEFNAETTFHADICTLSEKERQIKKOTALVELVKHKPKAT 540

QY 541 KEQLKAVMDDFAAFEVCKCKADDKETCFABEGKGLVAASQAALGL 585
DB 541 KEQLKAVMDDFAAFEVCKCKADDKETCFABEGKGLVAASQAALGL 585

RESULT 15
ABJ00986
ID ABJ00986 standard; protein; 585 AA.
XX

AC ABJ00986;
XX 05-SEP-2002 (first entry)
XX B lymphocyte stimulator protein binding peptide related protein.
XX
XX B lymphocyte stimulator protein binding protein; BlyS; immune disease;
XX allergy; proliferative disease; infectious disease; arteriosclerosis;
XX inflammatory disorder; hypergammaglobulinaemia; blood clotting;
XX ischaemia; graft-versus-host disease; neurodegenerative disease;
XX immunosuppressive; nephrotic; antirheumatic; antiarthritic;
XX neuroproliferative; cytostatic; immunostimulant; antitumour; anti-HIV;
XX antiasthmatic; antiallergic; thyromimetic; antianaemic; haemostatic;
XX dermatological; antiinflammatory; cardiant; ophthalmological; uropathic;
XX antidiabetic; antithyroid; antidepressant; hepatotropic.
XX
XX Homo sapiens.
XX WO200216411-A2.
XX 28-FEB-2002.
XX 17-AUG-2001; 2001WO-US025850.
XX 18-AUG-2000; 2000US-0226700P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Beltzer JP, Potter DM, Fleming TL, Rosen CA;
XX WPI; 2002-499775/53.
XX
XX The treatment of various diseases e.g. rheumatoid arthritis, comprises
XX administering B lymphocyte stimulator binding polypeptide.
XX
XX Disclosure; Page 379-382; 387pp; English.
XX
XX The present invention relates to the treatment, prevention or
XX amelioration of a disease or disorder associated with: aberrant B
XX lymphocyte stimulator (BlyS), BlyS receptor expression or activity; cells
XX of haematopoietic origin; or proliferative disease; and reducing,
XX inhibiting or stimulating immunoglobulin production, B cell proliferation
XX and graft rejection involving administration of BlyS binding polypeptide.
XX The BlyS binding polypeptides are used in the treatment, prevention or
XX amelioration of diseases such as immune system diseases, proliferative
XX diseases, infectious diseases, arteriosclerosis, inflammatory disorders,
XX allergies, infectious diseases, arteriosclerosis, inflammatory disorders,
XX hypergammaglobulinaemia, blood clotting disorders, ischaemia, and
XX neurodegenerative diseases. The present sequence is a protein described
XX in the invention
XX
XX Sequence 585 AA;

Query Match 100.0%; Score 585; DB 5; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFKDLGSENFKALVIAFAQYLOQCFFEDHVKLVNEVTEFAKTCVADSEAE 60
DB 1 DAHKSVAHRFKDLGSENFKALVIAFAQYLOQCFFEDHVKLVNEVTEFAKTCVADSEAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLPLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLPLVRPEV 120

QY 121 DVNCTAFHNEETFLKKLYEIAARRHPFYAPPELLFFAKRYKAAPTECCQADKAACLLP 180
DB 121 DVNCTAFHNEETFLKKLYEIAARRHPFYAPPELLFFAKRYKAAPTECCQADKAACLLP 180

QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRFPPKAEFAEYVKLVTDLT 240
DB 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRFPPKAEFAEYVKLVTDLT 240

RESULT 15
ABJ00986
ID ABJ00986 standard; protein; 585 AA.
XX

QY 481 LVNRRPCFSALEVDVETVPKFEFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540
DB 481 LVNRRPCFSALEVDVETVPKFEFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540
QY 541 KEQLKAVMDDFAAVFEKCKCKADDDKTCFAEKGKLVAAQAALGL 585
DB 541 KEQLKAVMDDFAAVFEKCKCKADDDKTCFAEKGKLVAAQAALGL 585

RESULT 13
AAE08578
ID AAE08578 standard; protein; 585 AA.
AC AAE08578;
DT 19-NOV-2001 (first entry)
DE Human serum albumin (HSA).
XX Human; albumin; cancer; cell proliferation; drug screening; biopsy.
XX Homo sapiens.
XX OS
XX PN US6274305-B1.
XX PD 14-AUG-2001.
XX PF 19-DEC-1996; 96US-00769746.
XX PR 19-DEC-1996; 96US-00769746.
XX PA (TUFT) UNIV TUFTS.
XX PI Sonnenschein C, Soto AM;
XX DR WPI; 2001-540371/60.
XX NR N-PSDB; AAD11488.
XX PT Measuring human cell proliferation, useful in drug screening to determine
PT the potential for inhibiting cancer cell proliferation and for evaluating
PT biopsied tumors, comprises employing albumin-derived peptide.
XX
XX Claim 1; Fig 1; 20pp; English.
XX
XX The invention related to a method for testing cancer cells. The method is
XX useful for measuring human cancer cell proliferation, particularly for
XX determining the potential for inhibiting cancer cells proliferation using
XX albumin-derived peptides. The invention is also useful for drug screening
XX assays, as well as for evaluating biopsied tumors. The present sequence
XX is human serum albumin (HSA) related to the invention
SQ Sequence 585 AA;

Query Match 100.0%; Score 585; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
DB 1 DAHKEVAHRFKDLGEENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVRPEV 120
QY 121 DVMCTAFHDNESTFLKKVLYELAREHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 121 DVMCTAFHDNESTFLKKVLYELAREHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLOKFGGRAPKAWAVARLSORFFPKAFPAEVSKLVTDLTK 240
DB 181 KLDELDEGKASSAKQRLKASLOKFGGRAPKAWAVARLSORFFPKAFPAEVSKLVTDLTK 240

QY 241 VHTCCCHGDLECCADRADLAKYICENQDSISSKLKCECEKFLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLECCADRADLAKYICENQDSISSKLKCECEKFLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADPFVESKDVCKNYABAKDVLGMLFLEYEARHPDYSVLLLLRLAKTYETTLK 360
DB 301 DLPSLAADPFVESKDVCKNYABAKDVLGMLFLEYEARHPDYSVLLLLRLAKTYETTLK 360
QY 361 CAAADPHECYAKVDFEFKPLVBEPPQNLIKQNCLEFQELGEYKFQNALLVRYTKKVPQVST 420
DB 361 CAAADPHECYAKVDFEFKPLVBEPPQNLIKQNCLEFQELGEYKFQNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
DB 421 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
QY 481 LVNRRPCFSALEVDVETVPKFEFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540
DB 481 LVNRRPCFSALEVDVETVPKFEFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540
QY 541 KEQLKAVMDDFAAVFEKCKCKADDDKTCFAEKGKLVAAQAALGL 585
DB 541 KEQLKAVMDDFAAVFEKCKCKADDDKTCFAEKGKLVAAQAALGL 585

RESULT 14
AAU75220
ID AAU75220 standard; protein; 585 AA.
XX
XX AC AAU75220;
XX DT 21-MAY-2002 (first entry)
XX DE Mature form of human serum albumin (HSA or HA).
XX
XX Albumin fusion protein; therapeutic protein; immune disorder;
XX autoimmune disorder; blood-related disorder; hyperproliferative
XX renal disorder; cardiovascular disorder; respiratory disorder;
XX neurological disorder; endocrine disorder; reproductive system disorder;
XX gastrointestinal disorder; infectious disease; wound healing;
XX human serum albumin; HSA; HA.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Protein 1. 585
XX /label= Mature_HSA
XX
XX WO200179271-A1.
XX PD 25-OCT-2001.
XX
XX 12-APR-2001; 2001WO-US012009.
XX
XX 12-APR-2000; 2000US-0229358P.
XX 25-APR-2000; 2000US-0199384P.
XX 21-DEC-2000; 2000US-0256931P.
XX
XX (PRIN-) PRINCIPIA PHARM CORP.
XX (DELZ) DELTA BIOTECHNOLOGY LTD.
XX
XX Ballance DJ, Sleep D, Turner AJ, Sadeghi H, Prior CP;
XX WPI; 2002-179329/23.
XX N-PSDB; ABK13862.
XX
XX New albumin fusion proteins with extended shelf life, useful for treating
XX leukemia, warts, hepatitis, multiple sclerosis and AIDS, comprises
XX therapeutic protein fused to albumin.
XX
XX Claim 1; Fig 15; 338pp; English.
XX
XX The present invention relates to albumin fusion proteins comprising a
CC

QY 541 KEOLKAVNDDFAAFVEKCKCKADDDKTCFAEBGKKLVAASQAALGL 585
Db 541 KEOLKAVNDDFAAFVEKCKCKADDDKTCFAEBGKKLVAASQAALGL 585

RESULT 12
ID ABB79006
AC ABB79006; standard; protein; 585 AA.
XX 01-AUG-2002 (first entry)
XX Human mature albumin protein SEQ ID NO:18.

XX Human; growth hormone; hGH; albumin; human serum albumin; HSA;
XX albumin fusion protein; cytostatic; anorectic; immunosuppressive;
XX antidiabetic; antirheumatic; antiarthritic; psoriatic; cancer;
XX non-Hodgkin's lymphoma; obesity; transplant rejection; psoriasis;
XX type I diabetes mellitus; rheumatoid arthritis.
XX Homo sapiens.

XX Key Location/Qualifiers
FH Domain 1..194
FT /label= 1
FT Domain 1..105
FT /label= subdomain
FT Disulfide-bond 53..62
FT Disulfide-bond 75..91
FT Disulfide-bond 90..101
FT Region 106..119
FT /note= "flexible inter-subdomain linker region"
FT Domain 120..194
FT /label= subdomain
FT Disulfide-bond 124..169
FT Disulfide-bond 168..177
FT Domain 195..387
FT /label= 2
FT Domain 195..291
FT /label= subdomain
FT Disulfide-bond 245..253
FT Disulfide-bond 265..279
FT Disulfide-bond 278..289
FT Region 292..315
FT /note= "flexible inter-subdomain linker region"
FT Domain 316..387
FT /label= subdomain
FT Disulfide-bond 316..361
FT Disulfide-bond 360..369
FT Domain 388..585
FT /label= 3
FT Domain 388..491
FT /label= subdomain
FT Disulfide-bond 392..438
FT Disulfide-bond 437..448
FT Disulfide-bond 461..477
FT Disulfide-bond 476..487
FT Region 492..511
FT /note= "flexible inter-subdomain linker region"
FT Domain 512..585
FT /label= subdomain
FT Disulfide-bond 514..559
FT Disulfide-bond 558..567

XX WO200179442-A2.
XX 25-OCT-2001.
XX 12-APR-2001; 2001WO-US011850.
XX 12-APR-2000; 2000US-0229358P.
XX 25-APR-2000; 2000US-0199384P.

PR 21-DEC-2000; 2000US-0256931P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Haseltine WA;
XX WPI; 2001-611723/70.
XX N-PSDB; ABB79006.
XX New albumin fusion proteins, useful for treating diseases and disorders
XX such as cancer, comprise therapeutic protein fused to albumin.
XX Claim 1; Fig 11; 413pp; English.
XX The present invention describes an albumin fusion protein (I) comprising
XX a therapeutic protein: X and (a fragment or variant of) albumin
XX comprising a the fully defined sequence in ABB79006 of 585 amino acids,
XX (where the fragment or variant has albumin or therapeutic protein: X
XX activity). (I) can have cytostatic, anorectic, immunosuppressive,
XX antidiabetic, antirheumatic, antiarthritic and psoriatic activities.
XX Albumin fusion proteins are stabilised therapeutic proteins e.g.
XX antibodies to C5, C242 and CD80 useful for treating various diseases and
XX disorders such as non-Hodgkin's lymphoma, cancer, obesity, transplant
XX rejection, type I diabetes mellitus, rheumatoid arthritis and psoriasis.
XX Fusing albumin to therapeutic proteins stabilises the therapeutic
XX protein, extends the shelf life and retains the in vitro or in vivo
XX biological activity. It also reduces the need to formulate protein
XX solutions with large excesses of carrier proteins to prevent loss of
XX therapeutic proteins due to factors such as binding to the container. The
XX fusion proteins are easily dispensed with a simple formulation requiring
XX minimal post storage manipulation. The fusion of therapeutic proteins to
XX albumin confers stability in aqueous or other solution. The present
XX sequence represents the mature human albumin (HA) protein which is used
XX in the exemplification of the present invention

XX Sequence 585 AA;
SQ

Query Match 100.0%; Score 585; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAKSEVAHRFKDLGEENFKALVLIAPAQLQCFEDHVKLVNEVTEFAKTCVADESAE 60
Db 1 DAKSEVAHRFKDLGEENFKALVLIAPAQLQCFEDHVKLVNEVTEFAKTCVADESAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNDNPLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNDNPLRLVRPEV 120
QY 121 DVMCTAFHDNBEFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Db 121 DVMCTAFHDNBEFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELRLDEGKASSAKQRLKCAASLQKFGERAFAKAWAVARLSORFPKAFPAEVSCLVTLTK 240
Db 181 KLDELRLDEGKASSAKQRLKCAASLQKFGERAFAKAWAVARLSORFPKAFPAEVSCLVTLTK 240
QY 241 VHTCCGDLLECCADDDRADLAKYICENODSISSKLKCECKEPLLEKSHCIAEVENDEMPA 300
Db 241 VHTCCGDLLECCADDDRADLAKYICENODSISSKLKCECKEPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDYCKNYAEAKDVLGWFYVEARRHPDYSVVLLRLAKTYTTLEKC 360
Db 301 DLPSLAADFVESKDYCKNYAEAKDVLGWFYVEARRHPDYSVVLLRLAKTYTTLEKC 360
QY 361 CAAADPEHCYAKVDFEFPKLVPEEPONLIKQCELFQELGEYKFNALLVRYTKKVPQVST 420
Db 361 CAAADPEHCYAKVDFEFPKLVPEEPONLIKQCELFQELGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRNLGVGSKCKCKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSRVTKCCTES 480
Db 421 PTLVEVSRNLGVGSKCKCKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSRVTKCCTES 480


```
QY 241 VHTCCGGDLLECCADRADLAKYICENQDSISSKLEKCEKELLESKSHCIAEVENDEMPA 300
Db 241 VHTCCGGDLLECCADRADLAKYICENQDSISSKLEKCEKELLESKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMLFYEARHPDYSVWLLRLAKTYETTLK 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMLFYEARHPDYSVWLLRLAKTYETTLK 360
QY 361 CAADPHCEYAKVDFEKFPLVEBPQNLIKONCELPQGLGEYKFNALLVRYTKVPQVST 420
Db 361 CAADPHCEYAKVDFEKFPLVEBPQNLIKONCELPQGLGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPFCFSALEVDVTVPKSFNAETFTPHADICTLSEKERQIKKOTALVELVHKPKAT 540
Db 481 LVNRRPFCFSALEVDVTVPKSFNAETFTPHADICTLSEKERQIKKOTALVELVHKPKAT 540
QY 541 KEQLKAVMDPFAAFVEKCKKADDKETCFABEGKGLVAASQAALGL 585
Db 541 KEQLKAVMDPFAAFVEKCKKADDKETCFABEGKGLVAASQAALGL 585

RESULT 10
AAE13129
ID AAE13129 standard; protein; 585 AA.
AC AAE13129;
DT 28-JAN-2002 (first entry)
DE Human albumin (HA).
KW Human; albumin; HA; fusion protein; therapeutic protein; vulnary;
KW immune system disorder; transplant rejection; blood related disorder;
KW myocardial infarction; hyperproliferative disorder; glomerulonephritis;
KW childhood acute myeloid leukemia; cardiovascular disorder; arrhythmia;
KW respiratory disorder; gene therapy; non-allergic rhinitis; noctropic;
KW neurological disease; Alzheimer's disease; reproductive system disorder;
KW endocrine disorder; pheochromocytoma; infectious disease; antiarthritic;
KW measles; gastrointestinal disorder; irritable bowel syndrome; syphilis;
KW wound healing; antiinflammatory; immunosuppressive; neuroprotective;
KW cardiac; cytostatic; antileukemic; antineumatic; antimicrobial;
KW renal disorder.
XX Homo sapiens.
OS
FH Key
FH Domain
FT 54..61 Location/Qualifiers
FT /label= Loop_I
FT 76..89 /label= Loop_II
FT 92..100 /label= Loop_III
FT 170..176 /label= Loop_IV
FT 247..252 /label= Loop_V
FT 266..277 /label= Loop_VI
FT 280..288 /label= Loop_VII
FT 362..368 /label= Loop_VIII
FT 439..447 /label= Loop_IX
FT 461..475 /label= Loop_X
FT 478..486 /label= Loop_XI
FT 560..566
```

```
FT 25-OCT-2001. /label= Loop_XII
FN WO200179443-A2.
PD 12-APR-2001; 2001WO-US011924.
XX 12-APR-2000; 2000US-0229358P.
XX 12-APR-2000; 2000US-0199384P.
XX 21-DEC-2000; 2000US-0256931P.
XX (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Haseltine WA;
PI N-PSDB; AAD21638.
DR WPI; 2001-616754/71.
DR N-PSDB; AAD21638.
XX Albumin fusion proteins comprising a therapeutic protein and albumin,
FT useful in the treating immune system disorders (e.g. transplant
FT rejection), blood related disorders (e.g. myocardial infarction) and
FT hyperproliferative disorders.
XX
PS Claim 1; Fig 9; 380pp; English.
XX
CC The invention relates to albumin fusion proteins comprising therapeutic
CC protein and human albumin (HA). Therapeutic proteins fused to albumin have
CC an extended shelf-life. The albumin fusion proteins are useful in the
CC treatment, prevention, diagnosis and/or detection of diseases, disorders
CC such as immune system disorders (e.g. transplant rejection), blood
CC related disorders (e.g. myocardial infarction), hyperproliferative
CC disorders (e.g. childhood acute myeloid leukemia), renal disorders (e.g.
CC glomerulonephritis), cardiovascular disorders (e.g. arrhythmias),
CC respiratory disorders (e.g. non-allergic rhinitis), neurological diseases
CC (e.g. Alzheimer's disease), endocrine disorders (e.g. pheochromocytoma),
CC reproductive system disorders (e.g. syphilis), infectious diseases (e.g.
CC measles), gastrointestinal disorders (e.g. irritable bowel syndrome), and
CC wound healing. Nucleic acids encoding albumin fusion protein is used in
CC gene therapy. The present sequence is human albumin (HA) protein
XX
SQ Sequence 585 AA;
Query Match 100.0%; Score 585; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHREFKDLGEENFKALVLIAPQYLOQCPPEHVKLVNEVTEFAKTCVADESAB 60
Db 1 DAHKSEVAHREFKDLGEENFKALVLIAPQYLOQCPPEHVKLVNEVTEFAKTCVADESAB 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVRREV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVRREV 120
QY 121 DVMCTAFHDNEETELFKYLYEIAARRHPYFYAPLLEFFAKRYKAAFTCCQAAADKAACLLP 180
Db 121 DVMCTAFHDNEETELFKYLYEIAARRHPYFYAPLLEFFAKRYKAAFTCCQAAADKAACLLP 180
QY 181 KLDELRLDEGKASSAKQRLKCAASLOKSGERAFKAWARLSQRPFAEFAEVSCLVTDLTK 240
Db 181 KLDELRLDEGKASSAKQRLKCAASLOKSGERAFKAWARLSQRPFAEFAEVSCLVTDLTK 240
QY 241 VHTCCGGDLLECCADRADLAKYICENQDSISSKLEKCEKELLESKSHCIAEVENDEMPA 300
Db 241 VHTCCGGDLLECCADRADLAKYICENQDSISSKLEKCEKELLESKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMLFYEARHPDYSVWLLRLAKTYETTLK 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMLFYEARHPDYSVWLLRLAKTYETTLK 360
QY 361 CAADPHCEYAKVDFEKFPLVEBPQNLIKONCELPQGLGEYKFNALLVRYTKVPQVST 420
Db 361 CAADPHCEYAKVDFEKFPLVEBPQNLIKONCELPQGLGEYKFNALLVRYTKVPQVST 420
```

121	QY	DB		DVMCTAFHDNEETFLKYLVEIARRHPYFYAPELFFPAKRYKAAETECQCAADKAACLLP	180
181	QY	DB		KLDELREBGGKASSAKORLKCASIQKFGERAFKAWAVARLSQRPFAEFAVSKLVTDLTJK	240
181	QY	DB		KLDELREBGGKASSAKORLKCASIQKFGERAFKAWAVARLSQRPFAEFAVSKLVTDLTJK	240
241	QY	DB		VHTECCCHGDLLECADRADLAKYICENQDSSISKLKECCKEPFLLEKSHHCIAEVENDEMPA	300
241	QY	DB		VHTECCCHGDLLECADRADLAKYICENQDSSISKLKECCKEPFLLEKSHHCIAEVENDEMPA	300
301	QY	DB		DLPSLAADFVESKDVCNKYAEAKDVFGLMFLYFYARRHPDYSVWLRLRLAKYETTLBKC	360
301	QY	DB		DLPSLAADFVESKDVCNKYAEAKDVFGLMFLYFYARRHPDYSVWLRLRLAKYETTLBKC	360
361	QY	DB		CAAADPHECYAKVDEPKPLVEBPQNLIKONCELFQOLGEYKFNQALVRYTKYQPVQST	420
361	QY	DB		CAAADPHECYAKVDEPKPLVEBPQNLIKONCELFQOLGEYKFNQALVRYTKYQPVQST	420
421	QY	DB		PTLVEVERNLKGVSCKCKHPEAKRMPCAEDYLSVNLGVHEKTPVSDRVTAKCCTES	480
421	QY	DB		PTLVEVERNLKGVSCKCKHPEAKRMPCAEDYLSVNLGVHEKTPVSDRVTAKCCTES	480
481	QY	DB		LVNRRPFCFSALVEVDYTPVPKEFNAETTFPHADICTLSEKERQIKKOTALVELVGHKPKAT	540
481	QY	DB		LVNRRPFCFSALVEVDYTPVPKEFNAETTFPHADICTLSEKERQIKKOTALVELVGHKPKAT	540
541	QY	DB		KEQLKAVNMDDPAAVFEVKCCKADDKETCFABEGKGLVAASQAALGL	585
541	QY	DB		KEQLKAVNMDDPAAVFEVKCCKADDKETCFABEGKGLVAASQAALGL	585

	XX	The invention relates to human albumin (HA) fusion proteins and their corresponding nucleic acid sequences. Therapeutic proteins fused to albumin or its fragments have an extended shelf-life. The albumin fusion proteins are useful in the treatment, prevention, diagnosis, and/or detection of diseases, disorders such as immune system disorders (e.g. transplant rejection), blood related disorders (e.g. myocardial infarction), hyperproliferative disorders (e.g. childhood acute myeloid leukaemia, metastatic renal cell carcinoma), metastatic melanoma, malignant melanoma, renal cell carcinoma), renal disorders (e.g. glomerulonephritis), cardiovascular disorders (e.g. arrhythmias), respiratory disorders (e.g. non-allergic rhinitis), neurological diseases (e.g. Alzheimer's disease), endocrine disorders (e.g. pheochromocytoma), reproductive system disorders (e.g. syphilis), infectious diseases (e.g. measles), gastrointestinal disorders (e.g. irritable bowel syndrome), HIV (human immunodeficiency virus) infection and wound healing. Nucleic acids encoding albumin fusion protein is used in gene therapy. The present sequence is human albumin
	SQ	Sequence 585 AA;

[illegible]

PR 17-JUN-1998; 98CN-00102506.
XX (HAIJ-) HAIJI BIOENGINEERING CO LTD.
PA
PI Li S, Lu D;
XX
XX WPI; 2000-351198/31.
DR N-PSDB; AAA10091.
XX
XX Process for preparing recombinant human serum albumin comprising yeast
PT biased sex codons - uses a recombinant DNA technique.
PT
PS Disclosure; Fig 1; 44pp; Chinese.
XX
XX The method relates to a method of recombinantly producing human serum
CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise
CC a yeast codon bias. The complete HSA gene (AAA10091) was generated as
CC three synthetic fragments (AAA10092-A10094) joined by recombinant DNA
CC technology. Each HSA fragment was synthesised from overlapping
CC oligonucleotide fragments that were extended. This sequence represents
CC the complete sequence of the HSA encoded by the human gene with a yeast
CC codon bias. The invention also covers a recombinant expression vector,
CC yeast host cells carrying the recombinant expression vector and the
CC process for producing human serum albumin in the yeast host cell,
CC especially in secretory mode
XX
XX Sequence 585 AA;
XX
Query Match 100.0%; Score 585; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRFDLGEENFKALVLI AFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHKSEVAHRFDLGEENFKALVLI AFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPLRLVPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPLRLVPEV 120
QY 121 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
Db 121 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
QY 181 KLDELDEGKASSAKORLKCASLOKGERAFKAWAVARLSQRFPAEFAEYKLVTLDTK 240
Db 181 KLDELDEGKASSAKORLKCASLOKGERAFKAWAVARLSQRFPAEFAEYKLVTLDTK 240
QY 241 VHTECCHGDLLECADRDADLAKYICENQDSISSKLKECEKPLEKSHCIAEVNDEMPPA 300
Db 241 VHTECCHGDLLECADRDADLAKYICENQDSISSKLKECEKPLEKSHCIAEVNDEMPPA 300
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYIYARRHPDYSVLLLRILAKTYETTLK 360
Db 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYIYARRHPDYSVLLLRILAKTYETTLK 360
QY 361 CAADAPHECVAKVDFBPKLVEBPQNLIKQNCLEFQELGEYKFQALLVRYTKVPQVST 420
Db 361 CAADAPHECVAKVDFBPKLVEBPQNLIKQNCLEFQELGEYKFQALLVRYTKVPQVST 420
QY 421 PTLVVEVSRNLKGVGSKCKHPEAKRMPCAEADYLSVVNLQNLVHLKTPVSRVTKCTTES 480
Db 421 PTLVVEVSRNLKGVGSKCKHPEAKRMPCAEADYLSVVNLQNLVHLKTPVSRVTKCTTES 480
QY 481 LVNRRPCFSALEVDVETVPKFEAEFTTHADICTLSEKERQIKQTALVELVKKPKAT 540
Db 481 LVNRRPCFSALEVDVETVPKFEAEFTTHADICTLSEKERQIKQTALVELVKKPKAT 540
QY 541 KEQLKAVMDFFAFAVEKCKKADDKETCFABEGKKUVAASQAALGL 585
Db 541 KEQLKAVMDFFAFAVEKCKKADDKETCFABEGKKUVAASQAALGL 585

RESULT 8
AAM52567
ID AAM52567 standard; protein; 585 AA.
XX
XX AAM52567;
XX
XX 05-FEB-2002 (first entry)
XX
XX Mature human serum albumin.
XX
XX Human; serum albumin; HA; antiinflammatory; immunosuppressive; cardiant;
KW nontropic; neuroprotective; gene therapy; immune disorder; wound healing;
KW hyperproliferative disorder; renal disorder; cardiovascular disorder;
KW respiratory disorder; neurological disease; endocrine disorder;
KW reproductive system disorder; infectious disease;
KW gastrointestinal disorder.
XX
XX Homo sapiens.
XX WO200179444-A2.
XX
XX 25-OCT-2001.
XX
XX 12-APR-2001; 2001WO-US012013.
XX
XX 12-APR-2000; 2000US-0229358P.
XX
XX 25-APR-2000; 2000US-0199384P.
XX
XX 21-DEC-2000; 2000US-0256931P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Haseltine WA;
PI
XX WPI; 2001-616755/71.
DR N-PSDB; ABA03057.
XX
XX Albumin fusion proteins comprising a therapeutic protein and albumin,
PT useful in the treating immune system disorders (e.g. transplant
PT rejection), blood related disorders (e.g. myocardial infarction) and
PT hyperproliferative disorders.
XX
XX Claim 1; Fig 15; 606pp; English.
XX
XX The present invention relates to albumin fusion proteins, which comprise
CC a therapeutic protein and albumin. The present sequence is the protein
CC sequence for mature human serum albumin (HA), which was used to generate
CC the fusion proteins of the present invention. The albumin fusion proteins
CC are useful in the treatment, prevention, diagnosis, and/or detection of
CC diseases/disorders such as immune system disorders (e.g. transplant
CC rejection), blood related disorders (e.g. childhood acute myeloid leukemia),
CC hyperproliferative disorders (e.g. glomerulonephritis), cardiovascular disorders (e.g.
CC arrhythmias), respiratory disorders (e.g. non-allergic rhinitis),
CC neurological diseases (e.g. Alzheimer's disease), endocrine disorders
CC (e.g. pheochromocytoma), reproductive system disorders (e.g. syphilis),
CC infectious diseases (e.g. measles), gastrointestinal disorders (e.g.
CC irritable bowel syndrome) and wound healing
XX
XX Sequence 585 AA;
XX
Query Match 100.0%; Score 585; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRFDLGEENFKALVLI AFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHKSEVAHRFDLGEENFKALVLI AFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPLRLVPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPLRLVPEV 120
QY 121 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180

QY 361 CAADPHCYAKVDEBKPLVEBPONLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420
DB 361 CAADPHCYAKVDEBKPLVEBPONLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETVYKPEFNAETTFHADICTLSEKERQIKKQATLVELVGHKPKAT 540
DB 481 LVNRRPCFSALEVDETVYKPEFNAETTFHADICTLSEKERQIKKQATLVELVGHKPKAT 540
QY 541 KEQLKAYMDDFAAFVEKCCCKADDDKTCFAEEGKKLVAASQAALGL 585
DB 541 KEQLKAYMDDFAAFVEKCCCKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 6

AY84873
ID AAY84873 standard; protein, 585 AA.
XX AC AAY84873;
XX DT 08-AUG-2000 (first entry)
XX DE Amino acid sequence of a human albumin protein.
XX KW Human; albumin; ischemic state; serum protein; metal ion salt;
KW peroperative ischemia; ischemia; myocardial infarction;
KW progressive coronary artery disease.
XX OS Homo sapiens.

Key Location/Qualifiers
FH Modified-site 1
FT /note= "optionally acetylated, and claimed under claim
FT 56"
XX WO200020840-A1.
XX 13-APR-2000.
XX 01-OCT-1999; 99WO-US022905.
XX 02-OCT-1999; 98US-00165591.
PR 02-OCT-1999; 98US-00165926.
PR 02-OCT-1999; 98US-0102738P.
PR 11-JAN-1999; 99US-0115392P.
XX (ISCH-) ISCHEMIA TECHNOLOGIES INC.
XX Bar-Or D, Lau E, Winkler JV;
XX WPI; 2000-303843/26.

PT New method for the continuous detection of ischemic states comprises
PT detecting and quantifying the existence of an alteration of the serum
PT protein albumin.
XX Disclosure; Page 97-100; 105pp; English.
XX The present sequence represents human albumin protein. The specification
XX describes a method for the continuous detection of ischemic states. The
CC method comprises detecting and quantifying the existence of an alteration
CC of the serum protein albumin. The method comprises contacting a
CC biological sample containing albumin from the patient with an excess
CC quantity of a metal ion salt, where the metal ion binds to the N-terminus
CC of naturally occurring human albumin, to form a mixture containing bound
CC metal ions and unbound metal ions, and then determining the amount of
CC metal ions bound to the albumin N-terminus. The amount of bound metal
CC ions is correlated to a known value to determine the occurrence or non-
CC occurrence of an ischemic event. The methods are useful for detection of

CC ischemic states. The methods are also useful for distinguishing
CC peroperative ischemia from ischemia caused by , amongst other things,
CC myocardial infarctions and progressive coronary artery disease
XX
SQ Sequence 585 AA;
Query Match 100.0%; Score 585; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHRFKDLGSENFKALVLTAFQYLOCCPEDHVKLVNEVTEFAKTCVADESAE 60
DB 1 DAHSEVAHRFKDLGSENFKALVLTAFQYLOCCPEDHVKLVNEVTEFAKTCVADESAE 60
QY 61 NCDKSLHTLFGDKLCTVATILRETYGEMADCCAKQBPNECECFLOHKDDNPNLRLVRREV 120
DB 61 NCDKSLHTLFGDKLCTVATILRETYGEMADCCAKQBPNECECFLOHKDDNPNLRLVRREV 120
QY 121 DVNCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
DB 121 DVNCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
QY 181 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
DB 181 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
QY 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKKECEKPLELLEKSHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECCADDRADLAKYICENQDSISSKKECEKPLELLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLVEYARRHPDYVSVLLRLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLVEYARRHPDYVSVLLRLAKTYETTLK 360
QY 361 CAADPHCYAKVDEBKPLVEBPONLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420
DB 361 CAADPHCYAKVDEBKPLVEBPONLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETVYKPEFNAETTFHADICTLSEKERQIKKQATLVELVGHKPKAT 540
DB 481 LVNRRPCFSALEVDETVYKPEFNAETTFHADICTLSEKERQIKKQATLVELVGHKPKAT 540
QY 541 KEQLKAYMDDFAAFVEKCCCKADDDKTCFAEEGKKLVAASQAALGL 585
DB 541 KEQLKAYMDDFAAFVEKCCCKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 7

AY83946
ID AAY83946 standard; protein, 585 AA.
XX AC AAY83946;
XX DT 28-JUL-2000 (first entry)
XX DE Yeast codon-biased recombinant human serum albumin protein.
XX KW Recombinant; human serum albumin; HSA; yeast codon bias; host cell;
KW overlapping oligonucleotide; expression vector.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN CN1239103-A.
XX 22-DEC-1999.
XX 17-JUN-1998; 98CN-00102506.
XX

PN W09523857-A1.
 XX 08-SEP-1995.
 XX 01-MAR-1995; 95WO-GB000434.
 XX 05-MAR-1994; 94GB-00004270.
 XX (DELZ) DELTA BIOTECHNOLOGY LTD.
 XX KerryWilliams SM, Gilbert SC;
 XX WPI: 1995-320572/41.
 XX N-PSDB; AAQ98695.
 XX Yeast with reduced levels of aspartyl protease 3 proteolytic activity -
 XX used to secrete human albumin without prodn. of the 45 kD fragment.
 XX Example 1; Page 26-28; 50pp; English.
 XX The cDNA given in AAQ98695, which encodes HSA (AAR80301), was subjected
 CC to site-directed mutagenesis to investigate the role of endoproteases in
 CC the generation of a 45 kDa albumin fragment obtd. when the cDNA is
 CC expressed in S. cerevisiae. Mutations were: R410A; L409V, V409A;
 CC and R410A, K413Q, K414Q. The latter set of mutations, especially,
 CC improved stability of HSA to yeast yap3p proteolytic cleavage, allowing
 CC increased prodn. of recombinant HSA. (Updated on 25-MAR-2003 to correct
 CC PI field.)
 XX FI field.)
 XX Sequence 585 AA;
 XX
 Query Match 100.0%; Score 585; DB 2; Length 585;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHSEVAHFRKDLGEENFKALVLIAPAQYLQOCFFEDHVKLVNEVTEFAKTCVADESAAE 60
 DB 1 DAHSEVAHFRKDLGEENFKALVLIAPAQYLQOCFFEDHVKLVNEVTEFAKTCVADESAAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVPEV 120
 QY 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
 DB 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
 QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEPAEVSKLVTDLTK 240
 DB 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEPAEVSKLVTDLTK 240
 QY 241 VHTECHGDLLECCADRADLAKYICENQDSISSKLKECEKPLEKSHCIAEVENDEMPA 300
 DB 241 VHTECHGDLLECCADRADLAKYICENQDSISSKLKECEKPLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYFYARRHPDYSVLLLLRLAKTYETTLK 360
 DB 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYFYARRHPDYSVLLLLRLAKTYETTLK 360
 QY 361 CAADPHCEYAKVDFEFPQNLIKQNCBELFQELGEYKFQALLVRYTKKVPQVST 420
 DB 361 CAADPHCEYAKVDFEFPQNLIKQNCBELFQELGEYKFQALLVRYTKKVPQVST 420
 QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 QY 481 LVNRRPFCSALEVDVETVPKFEAFNFTTHADICTLSEKEROIKQTALVELVGHKPKAT 540
 DB 481 LVNRRPFCSALEVDVETVPKFEAFNFTTHADICTLSEKEROIKQTALVELVGHKPKAT 540
 QY 541 KEQLKAVMDDFAAVFVECKCKADDKETCFABEGKGLVAASQAALGL 585

DB 541 KEQLKAVMDDFAAVFVECKCKADDKETCFABEGKGLVAASQAALGL 585
 RESULT 5
 AAO20111
 ID AAO20111 standard; protein; 585 AA.
 XX AC AAO20111;
 XX 06-AUG-2002 (first entry)
 XX HSA protein sequence related to the growth hormone protein.
 XX DE Serum albumin-growth hormone fusion protein; growth hormone;
 XX KW Down's syndrome.
 XX OS Unidentified.
 XX PN KR99076789-A.
 XX 15-OCT-1999.
 XX 25-JUN-1998; 98KR-00704914.
 XX 30-DEC-1995; 95GB-00026733.
 XX 19-DEC-1996; 96WO-GB003164.
 XX (DELZ) DELTA BIOTECHNOLOGY LTD.
 XX Ballance DJ;
 XX WPI: 1997-363680/33.
 XX N-PSDB; AAK99568.
 XX Serum albumin-growth hormone fusion protein - useful to treat growth
 XX hormone related diseases, e.g. Down's syndrome.
 XX Disclosure; Fig 6; 21pp; Korean.
 XX The invention relates to a serum albumin-growth hormone fusion protein -
 CC useful to treat growth hormone related diseases such as Down's syndrome.
 CC This sequence represents a HSA protein related to the serum albumin-
 CC growth hormone protein of the invention
 XX Sequence 585 AA;
 XX
 Query Match 100.0%; Score 585; DB 2; Length 585;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHSEVAHFRKDLGEENFKALVLIAPAQYLQOCFFEDHVKLVNEVTEFAKTCVADESAAE 60
 DB 1 DAHSEVAHFRKDLGEENFKALVLIAPAQYLQOCFFEDHVKLVNEVTEFAKTCVADESAAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVPEV 120
 QY 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
 DB 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
 QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEPAEVSKLVTDLTK 240
 DB 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEPAEVSKLVTDLTK 240
 QY 241 VHTECHGDLLECCADRADLAKYICENQDSISSKLKECEKPLEKSHCIAEVENDEMPA 300
 DB 241 VHTECHGDLLECCADRADLAKYICENQDSISSKLKECEKPLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYFYARRHPDYSVLLLLRLAKTYETTLK 360
 DB 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYFYARRHPDYSVLLLLRLAKTYETTLK 360

99 478 81.7 692 7 ADD06565 Human Ckb
100 470 80.3 585 2 AAR26362 Synthetic

ALIGNMENTS

RESULT 1
ID AAP90388 standard; protein; 585 AA.
XX AAP90388;
AC AAP90388;
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 01-NOV-1989 (first entry)
XX Mature human serum albumin polypeptide.
XX Human serum albumin; mature protein; new polypeptides; plasma expanders.
XX Homo sapiens; (Human).
XX EP322094-A.
XX 28-JUN-1989.
XX 25-OCT-1988; 88EP-00310000.
XX 30-OCT-1987; 87GB-00025529.
XX (DELZ) DELTA BIOTECHNOLOGY LTD.
XX Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;
XX WPI; 1989-196464/26.
XX N-PSDB; AAN90128.
XX New N-terminal fragments of human serum albumin - esp. useful as blood
XX plasma expanders.
XX Disclosure; Fig 2; 20pp; English.
XX Mature protein of human serum albumin (see corresp. AAN90128). Used to
XX make new N-terminal fragments which are used as plasma expanders, or as
XX substitutes for HSA or BSA, in tissue culture media. (Updated on 25-MAR-
XX 2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
XX field)
XX Sequence 585 AA;
Query Match 100.0%; Score 585; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKEVAHRFKDLGEENFKALVLIATAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHKEVAHRFKDLGEENFKALVLIATAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLRLVRPEV 120
Qy 121 DVNCTAFHNDNEETFLKXLYETARRHPYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 121 DVNCTAFHNDNEETFLKXLYETARRHPYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Qy 181 KLDELDEGKASSAKQRLKASIQFGERAFAKAWAVARLSORFFPKAEPAEVSKLVTDLTK 240
Db 181 KLDELDEGKASSAKQRLKASIQFGERAFAKAWAVARLSORFFPKAEPAEVSKLVTDLTK 240
Qy 241 VHTECCHGDLLECADRADLAKYICENQDSISSKULKECCEKPLEKSHCIAEVENDEMPA 300

Db 241 VHTECCHGDLLECADRADLAKYICENQDSISSKULKECCEKPLEKSHCIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCNKYAEAKDVFGLMFLYAYARRHPDYSVVLLRLAKTYETTLKXC 360
Db 301 DLPSLAADFVESKDVCNKYAEAKDVFGLMFLYAYARRHPDYSVVLLRLAKTYETTLKXC 360
Qy 361 CAAADPHECYAKVPDEFKPLVEEPQNLIKONCELPOLGEYKFNALLVRYTKKYPQVST 420
Db 361 CAAADPHECYAKVPDEFKPLVEEPQNLIKONCELPOLGEYKFNALLVRYTKKYPQVST 420
Qy 421 PTLVEVSRLNGLKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRLNGLKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Qy 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKQTALVELVKHKPKAT 540
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKQTALVELVKHKPKAT 540
Qy 541 KEQLKAVMDFFAAFEKCKKADDDKTCFAEBEGKLVAAASQAALGL 585
Db 541 KEQLKAVMDFFAAFEKCKKADDDKTCFAEBEGKLVAAASQAALGL 585

RESULT 2
AAR05318
ID AAR05318 standard; protein; 585 AA.
XX AAR05318;
AC AAR05318;
DT 08-OCT-1990 (first entry)
XX Human serum albumin gene product.
XX Human serum albumin; HSA-A; yeast; ds.
XX Homo sapiens.
XX JP02117384-A.
XX 01-MAY-1990.
XX 26-OCT-1988; 88JP-00268302.
XX 26-OCT-1988; 88JP-00268302.
XX (TOFU) TOA NENRYO KOGYO KK.
XX WPI; 1990-176228/23.
XX N-PSDB; AAQ04719.
XX Human serum albumin prepn. by yeast host - by culturing transformed
XX plasmid yeast to produce serum, and removing it.
XX Disclosure; Page ?; -pp; Japanese.
XX Mature HSA-A may be produced using the sequence incorporated into a
XX plasmid vector with suitable controllers, and transferred to a yeast
XX expression system
XX Sequence 585 AA;
Query Match 100.0%; Score 585; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKEVAHRFKDLGEENFKALVLIATAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHKEVAHRFKDLGEENFKALVLIATAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLRLVRPEV 120

OM protein - protein search, using sw model

Run on: April 19, 2004, 16:01:29 ; Search time 60 Seconds
(without alignments)
2754.838 Million cell updates/sec

Title: US-09-832-929-18

Perfect score: 585

Sequence: 1 DAHSEVAHFKOLGEENFK.....TCFAEGKKLVAASQAALGL 585

Scoring table: OLIGO 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 100 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003as.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	585	100.0	585	1	AAP90388 Mature hu
2	585	100.0	585	2	AAR05318 Human ser
3	585	100.0	585	2	AAR08457 Human ser
4	585	100.0	585	2	AAR80301 Human ser
5	585	100.0	585	2	AAO20111 HSA prote
6	585	100.0	585	3	AAV84873 Amino aci
7	585	100.0	585	3	AAV83946 Yeast cod
8	585	100.0	585	4	AAW52567 Mature hu
9	585	100.0	585	4	AAE12403 Human alb
10	585	100.0	585	4	AAE13129 Human alb
11	585	100.0	585	4	AAE13399 Human alb
12	585	100.0	585	4	ABW79006 Human mat
13	585	100.0	585	4	AAE08578 Human ser
14	585	100.0	585	5	AAU75220 Mature fo
15	585	100.0	585	5	ABJ00986 B lymphoc
16	585	100.0	585	5	ABG63321 Human ser
17	585	100.0	585	5	ABG33847 Human B L
18	585	100.0	585	5	ABG71291 Glycosyla
19	585	100.0	585	6	ABR55695 Human alb
20	585	100.0	585	7	ABR42606 Human ser
21	585	100.0	585	7	ADC16767 Human ser
22	585	100.0	585	7	ADD06469 Human ser
23	585	100.0	585	7	ADD68016 Mature fo
24	585	100.0	609	3	AAH36542 Recombina
25	585	100.0	609	3	AAV78147 Pre human

DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE CpeB (Fragment).
GN CPB.
OS uncultured Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=159733;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22753450; PubMed=12871235;
RA Steglich C., Post A.F., Hess W.R.;
RT "Analysis of natural populations of prochlorococcus spp. in the
RT northern Red Sea using phycoerythrin gene sequences.";
RL Environ. Microbiol. 5:681-690(2003).
DR EMBL; AF438692; AAP97619.1; -.
FT NON_TER 1
FT NON_TER 74
SQ SEQUENCE 74 AA; 8336 MW; 55AC49080B8CBC7A CRC64;

Query Match 1.2%; Score 7; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 ICENQDS 270
Db 1 ICENQDS 7

RESULT 75
Q7WVKO PRELIMINARY; PRT; 74 AA.
AC Q7WVKO;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE CpeB (Fragment).
GN CPB.
OS uncultured Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=159733;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22753450; PubMed=12871235;
RA Steglich C., Post A.F., Hess W.R.;
RT "Analysis of natural populations of prochlorococcus spp. in the
RT northern Red Sea using phycoerythrin gene sequences.";
RL Environ. Microbiol. 5:681-690(2003).
DR EMBL; AF438693; AAP97620.1; -.
FT NON_TER 1
FT NON_TER 74
SQ SEQUENCE 74 AA; 8322 MW; C09708BADB8D1B8F CRC64;

Query Match 1.2%; Score 7; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 264 ICENQDS 270
Db 1 ICENQDS 7

Search completed: April 19, 2004, 16:07:17
Job time : 48 secs

```

QY 264 ICNQDS 270
Db 1 ICNQDS 7

RESULT 70
Q7WVK5 PRELIMINARY; PRT; 74 AA.
AC Q7WVK5;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE CpeB (Fragment).
GN CPEB.
OS uncultured Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=159733;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22753450; PubMed=12871235;
RA Steglich C., Post A.F., Hess W.R.;
RT "Analysis of natural populations of Prochlorococcus spp. in the
RT northern Red Sea using phycoerythrin gene sequences.";
RL Environ. Microbiol. 5:681-690(2003).
DR EMBL; AF438686; AAP97613.1; -.
FT NON_TER 1
FT NON_TER 74
SQ SEQUENCE 74 AA; 8293 MW; 8EC5DD0DEB8C41F9 CRC64;

Query Match 1.2%; Score 7; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 ICNQDS 270
Db 1 ICNQDS 7

RESULT 71
Q7WVK4 PRELIMINARY; PRT; 74 AA.
AC Q7WVK4;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE CpeB (Fragment).
GN CPEB.
OS uncultured Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=159733;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22753450; PubMed=12871235;
RA Steglich C., Post A.F., Hess W.R.;
RT "Analysis of natural populations of Prochlorococcus spp. in the
RT northern Red Sea using phycoerythrin gene sequences.";
RL Environ. Microbiol. 5:681-690(2003).
DR EMBL; AF438688; AAP97615.1; -.
FT NON_TER 1
FT NON_TER 74
SQ SEQUENCE 74 AA; 8398 MW; 4EA460980B8CAEC3 CRC64;

Query Match 1.2%; Score 7; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 ICNQDS 270
Db 1 ICNQDS 7

RESULT 72
Q7WVK3 PRELIMINARY; PRT; 74 AA.
AC Q7WVK3;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE CpeB (Fragment).
GN CPEB.
OS uncultured Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=159733;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22753450; PubMed=12871235;
RA Steglich C., Post A.F., Hess W.R.;
RT "Analysis of natural populations of Prochlorococcus spp. in the
RT northern Red Sea using phycoerythrin gene sequences.";
RL Environ. Microbiol. 5:681-690(2003).
DR EMBL; AF438690; AAP97617.1; -.
FT NON_TER 1
FT NON_TER 74
SQ SEQUENCE 74 AA; 8327 MW; C125DD0DE1814683 CRC64;

Query Match 1.2%; Score 7; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 ICNQDS 270
Db 1 ICNQDS 7

RESULT 73
Q7WVK2 PRELIMINARY; PRT; 74 AA.
AC Q7WVK2;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE CpeB (Fragment).
GN CPEB.
OS uncultured Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=159733;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22753450; PubMed=12871235;
RA Steglich C., Post A.F., Hess W.R.;
RT "Analysis of natural populations of Prochlorococcus spp. in the
RT northern Red Sea using phycoerythrin gene sequences.";
RL Environ. Microbiol. 5:681-690(2003).
DR EMBL; AF438691; AAP97618.1; -.
FT NON_TER 1
FT NON_TER 74
SQ SEQUENCE 74 AA; 8349 MW; C8B5DD0DECLF2A8C CRC64;

Query Match 1.2%; Score 7; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 ICNQDS 270
Db 1 ICNQDS 7

RESULT 74
Q7WVK1 PRELIMINARY; PRT; 74 AA.
AC Q7WVK1;

```


Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGKV 433
Db 53 SRNLGKV 59

RESULT 66

Q9BX1 PRELIMINARY; PRT; 73 AA.

AC Q9BX1, 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE Genome polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=NIV-7;
RA Jha J., Arankalle V.A.;
RT "Phylogenetic Analysis of Indian HCV Isolates";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT, THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND RNA (BY SIMILARITY).
CC EMBL; AF134744; AAG09125.1;
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Polyprotein.
FT NON_TER 1
FT NON_TER 73
FT NON_TER 73
SQ SEQUENCE 73 AA; 8073 MW; A01A25878AA44D55 CRC64;

Query Match 1.2%; Score 7; DB 12; Length 73;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGKV 433
Db 53 SRNLGKV 59

RESULT 67

Q7WVK8 PRELIMINARY; PRT; 74 AA.

AC Q7WVK8, 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DE Genome polyprotein (Fragment).
OS uncultured Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=159733;
RN [1]
RN SEQUENCE FROM N.A.
RC MEDLINE=22753450; PubMed=12871235;
RA Steglich C., Post A.F., Hess W.R.;
RT "Analysis of natural populations of Prochlorococcus spp. in the
RT northern Red Sea using phycoerythrin gene sequences.";
RL Environ. Microbiol. 5:681-690(2003).
DR EMBL; AF438681; AAP97608.1;
FT NON_TER 1
FT NON_TER 74
FT NON_TER 74
SQ SEQUENCE 74 AA; 8293 MW; C53C4E6A3ACD498B CRC64;

Query Match 1.2%; Score 7; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

SQ SEQUENCE 74 AA; 8344 MW; D7B5949DEB995489 CRC64;

Query Match 1.2%; Score 7; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 ICENQDS 270
Db 1 ICENQDS 7

RESULT 68

Q7WVK7 PRELIMINARY; PRT; 74 AA.

AC Q7WVK7, 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DE Genome polyprotein (Fragment).
OS uncultured Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=159733;
RN [1]
RN SEQUENCE FROM N.A.
RC MEDLINE=22753450; PubMed=12871235;
RA Steglich C., Post A.F., Hess W.R.;
RT "Analysis of natural populations of Prochlorococcus spp. in the
RT northern Red Sea using phycoerythrin gene sequences.";
RL Environ. Microbiol. 5:681-690(2003).
DR EMBL; AF438682; AAP97609.1;
FT NON_TER 1
FT NON_TER 74
FT NON_TER 74
SQ SEQUENCE 74 AA; 8385 MW; D7A703F80B8CB76 CRC64;

Query Match 1.2%; Score 7; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 ICENQDS 270
Db 1 ICENQDS 7

RESULT 69

Q7WVK6 PRELIMINARY; PRT; 74 AA.

AC Q7WVK6, 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DE Genome polyprotein (Fragment).
OS uncultured Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=159733;
RN [1]
RN SEQUENCE FROM N.A.
RC MEDLINE=22753450; PubMed=12871235;
RA Steglich C., Post A.F., Hess W.R.;
RT "Analysis of natural populations of Prochlorococcus spp. in the
RT northern Red Sea using phycoerythrin gene sequences.";
RL Environ. Microbiol. 5:681-690(2003).
DR EMBL; AF438684; AAP97611.1;
FT NON_TER 1
FT NON_TER 74
FT NON_TER 74
SQ SEQUENCE 74 AA; 8293 MW; C53C4E6A3ACD498B CRC64;

Query Match 1.2%; Score 7; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; AF134748; AAG09129.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 73
SQ SEQUENCE 73 AA; 8335 MW; 1377463A2425A984 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 12; Length 73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGKV 433
DB 53 SRNLGKV 59

RESULT 63
Q9E8X6 PRELIMINARY; PRT; 73 AA.
ID Q9E8X6;
AC Q9E8X6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Genome polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RC STRAIN=NIV-1;
RA Jha J., Arankalle V.A.;
RT "Phylogenetic Analysis of Indian HCV Isolates.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; AF134739; AAG09120.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 73
SQ SEQUENCE 73 AA; 8157 MW; A41CAE64E0B3306A CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 12; Length 73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGKV 433
DB 53 SRNLGKV 59

RESULT 64
Q9E8W6 PRELIMINARY; PRT; 73 AA.
ID Q9E8W6;
AC Q9E8W6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
```

```
01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Genome polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RC STRAIN=NIV-15;
RA Jha J., Arankalle V.A.;
RT "Phylogenetic Analysis of Indian HCV Isolates.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; AF134749; AAG09130.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 73
SQ SEQUENCE 73 AA; 8173 MW; 0EA0DE64EE261999 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 12; Length 73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGKV 433
DB 53 SRNLGKV 59

RESULT 65
Q9E8W8 PRELIMINARY; PRT; 73 AA.
ID Q9E8W8;
AC Q9E8W8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Genome polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RC STRAIN=NIV-13;
RA Jha J., Arankalle V.A.;
RT "Phylogenetic Analysis of Indian HCV Isolates.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; AF134747; AAG09128.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 73
SQ SEQUENCE 73 AA; 8182 MW; 1A560ECE8E8112CB CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 12; Length 73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
ID Q9B8W9 PRELIMINARY; PRT; 73 AA.
AC Q9B8W9;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Genome polyprotein (Fragment).
DE Hepatitis C virus.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIV-12;
RA Jha J., Arankalle V.A.;
RT "Phylogenetic Analysis of Indian HCV Isolates.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; AF134746; AAC09127.1; -
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Polyprotein.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 73
SQ SEQUENCE 73 AA; 8266 MW; F01CF65442F8F648 CRC64;

Query Match 1.2%; Score 7; DB 12; Length 73;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGV 433
Db 53 SRNLGV 59

RESULT 60
Q9B8X4 PRELIMINARY; PRT; 73 AA.
AC Q9B8X4;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Genome polyprotein (Fragment).
DE Hepatitis C virus.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIV-3;
RA Jha J., Arankalle V.A.;
RT "Phylogenetic Analysis of Indian HCV Isolates.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; AF134741; AAC09122.1; -
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Polyprotein.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 73
SQ SEQUENCE 73 AA; 8266 MW; F01CF65442F8F648 CRC64;

Query Match 1.2%; Score 7; DB 12; Length 73;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGV 433
Db 53 SRNLGV 59

RESULT 61
Q9B8X0 PRELIMINARY; PRT; 73 AA.
AC Q9B8X0;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Genome polyprotein (Fragment).
DE Hepatitis C virus.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIV-11;
RA Jha J., Arankalle V.A.;
RT "Phylogenetic Analysis of Indian HCV Isolates.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; AF134745; AAC09126.1; -
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Polyprotein.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 73
SQ SEQUENCE 73 AA; 8165 MW; 6307F954E0AE14D1 CRC64;

Query Match 1.2%; Score 7; DB 12; Length 73;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGV 433
Db 53 SRNLGV 59

RESULT 62
Q9B8W7 PRELIMINARY; PRT; 73 AA.
AC Q9B8W7;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Genome polyprotein (Fragment).
DE Hepatitis C virus.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIV-14;
RA Jha J., Arankalle V.A.;
RT "Phylogenetic Analysis of Indian HCV Isolates.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
```

```
SQ SEQUENCE 73 AA; 8144 MW; F01D585786842CAE CRC64;

Query Match 1.2%; Score 7; DB 12; Length 73;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGV 433
Db 53 SRNLGV 59

RESULT 61
Q9B8X0 PRELIMINARY; PRT; 73 AA.
AC Q9B8X0;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Genome polyprotein (Fragment).
DE Hepatitis C virus.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIV-11;
RA Jha J., Arankalle V.A.;
RT "Phylogenetic Analysis of Indian HCV Isolates.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; AF134745; AAC09126.1; -
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Polyprotein.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 73
SQ SEQUENCE 73 AA; 8165 MW; 6307F954E0AE14D1 CRC64;

Query Match 1.2%; Score 7; DB 12; Length 73;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGV 433
Db 53 SRNLGV 59

RESULT 62
Q9B8W7 PRELIMINARY; PRT; 73 AA.
AC Q9B8W7;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Genome polyprotein (Fragment).
DE Hepatitis C virus.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIV-14;
RA Jha J., Arankalle V.A.;
RT "Phylogenetic Analysis of Indian HCV Isolates.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
```

```
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL; AF003008; BAB52462.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 71 AA; 7745 MW; 47899ACB6E04C805 CRC64;

Query Match 1.2%; Score 7; DB 16; Length 71;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 288 HCAIEVE 294
DB 37 HCAIEVE 43

RESULT 56
Q9E8X2 PRELIMINARY; PRT; 73 AA.
AC Q9E8X2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Genome polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIV-5;
RA Jha J., Arankalle V.A.;
RT "Phylogenetic Analysis of Indian HCV Isolates.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; AF134743; AAG09124.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Polyprotein.
FT NON_TER 1 73
FT NON_TER 73 73
SQ SEQUENCE 73 AA; 8208 MW; E92D59FE0AB4A64 CRC64;

Query Match 1.2%; Score 7; DB 12; Length 73;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGV 433
DB 53 SRNLGV 59

RESULT 57
Q9E8X3 PRELIMINARY; PRT; 73 AA.
AC Q9E8X3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Genome polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIV-4;
RA Jha J., Arankalle V.A.;
RT "Phylogenetic Analysis of Indian HCV Isolates.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; AF134740; AAG09121.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Polyprotein.
FT NON_TER 1 73
FT NON_TER 73 73
SQ SEQUENCE 73 AA; 8208 MW; E92D59FE0AB4A64 CRC64;

Query Match 1.2%; Score 7; DB 12; Length 73;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGV 433
DB 53 SRNLGV 59
```

```
RA Jha J., Arankalle V.A.;
RT "Phylogenetic Analysis of Indian HCV Isolates.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; AF134742; AAG09123.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Polyprotein.
FT NON_TER 1 73
FT NON_TER 73 73
SQ SEQUENCE 73 AA; 8161 MW; 0AF5967DE80A07E9 CRC64;

Query Match 1.2%; Score 7; DB 12; Length 73;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGV 433
DB 53 SRNLGV 59

RESULT 58
Q9E8X5 PRELIMINARY; PRT; 73 AA.
AC Q9E8X5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Genome polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIV-2;
RA Jha J., Arankalle V.A.;
RT "Phylogenetic Analysis of Indian HCV Isolates.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; AF134740; AAG09121.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Polyprotein.
FT NON_TER 1 73
FT NON_TER 73 73
SQ SEQUENCE 73 AA; 8203 MW; AC6AD14DE0A66021 CRC64;

Query Match 1.2%; Score 7; DB 12; Length 73;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGV 433
DB 53 SRNLGV 59

RESULT 59
Q9E8W9
```

DE Genome polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTX;
RA Mueller H.M., Selzer S., Heller A.E., Pfaff E., Theilmann L.;
RT "Recurrent hepatitis C virus infection after liver transplantation in
a patient negative for viral proteins: A possible role of immune
tolerance".
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; U45471; AAA6927.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
KW Polyprotein.
FT NON_TER 1 1
FT NON_TER 65 65
FT SEQUENCE 65 AA; 7335 MW; A36A6E786472DE28 CRC64;
SQ
Query Match 1.2%; Score 7; DB 12; Length 65;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 427 SRNLGKV 433
Db 51 SRNLGKV 57
RESULT 53
Q68526 PRELIMINARY; PRT; 65 AA.
AC Q68526
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Genome polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTX;
RA Mueller H.M., Selzer S., Heller A.E., Pfaff E., Theilmann L.;
RT "Recurrent hepatitis C virus infection after liver transplantation in
a patient negative for viral proteins: A possible role of immune
tolerance".
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; U45469; AAA6925.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
KW Polyprotein.
FT NON_TER 1 1
FT NON_TER 65 65
FT SEQUENCE 65 AA; 7335 MW; A36A6E786472DE28 CRC64;
SQ

Query Match 1.2%; Score 7; DB 12; Length 65;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 427 SRNLGKV 433
Db 51 SRNLGKV 57
RESULT 54
O40646 PRELIMINARY; PRT; 67 AA.
ID O40646
AC O40646;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Core protein (Genome polyprotein) (fragment).
OS Hepatitis C virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=40271;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cap1;
RA Dettori S., Spada E., Rapicetta M.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; Y09368; CAA70345.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR Pfam; PF01543; HCV_capsid; 1.
KW Polyprotein.
FT NON_TER 1 1
FT NON_TER 67 67
FT SEQUENCE 67 AA; 7634 MW; C47EC465C2B506AB CRC64;
SQ
Query Match 1.2%; Score 7; DB 12; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 427 SRNLGKV 433
Db 60 SRNLGKV 66
RESULT 55
Q98A75 PRELIMINARY; PRT; 71 AA.
ID Q98A75
AC Q98A75;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein ms16120.
GN MS16120.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RL "Complete genome structure of the nitrogen-fixing symbiotic bacterium

```
Db 51 SRNLGKV 57
|||||
RESULT 49
ID Q68531 PRELIMINARY; PRT; 65 AA.
AC Q68531;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Genome polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTx;
RA Mueller H.M., Selzer S., Heller A.E., Pfaff E., Theilmann L.;
RT "Recurrent hepatitis C virus infection after liver transplantation in
RT a patient negative for viral proteins: A possible role of immune
RT tolerance."
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -|- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; U45474; AAA86930.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Polyprotein.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 65
SQ SEQUENCE 65 AA; 7335 MW; A36A6E786472DE28 CRC64;

Query Match 1.2%; Score 7; DB 12; Length 65;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGKV 433
|||||
Db 51 SRNLGKV 57

RESULT 50
ID Q68532 PRELIMINARY; PRT; 65 AA.
AC Q68532;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Genome polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTx;
RA Mueller H.M., Selzer S., Heller A.E., Pfaff E., Theilmann L.;
RT "Recurrent hepatitis C virus infection after liver transplantation in
RT a patient negative for viral proteins: A possible role of immune
RT tolerance."
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -|- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; U45474; AAA86930.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Polyprotein.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 65
SQ SEQUENCE 65 AA; 7335 MW; A36A6E786472DE28 CRC64;

Query Match 1.2%; Score 7; DB 12; Length 65;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGKV 433
|||||
Db 51 SRNLGKV 57

RESULT 51
ID Q68527 PRELIMINARY; PRT; 65 AA.
AC Q68527;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Genome polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTx;
RA Mueller H.M., Selzer S., Heller A.E., Pfaff E., Theilmann L.;
RT "Recurrent hepatitis C virus infection after liver transplantation in
RT a patient negative for viral proteins: A possible role of immune
RT tolerance."
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -|- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; U45470; AAA86926.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Polyprotein.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 65
SQ SEQUENCE 65 AA; 7335 MW; A36A6E786472DE28 CRC64;

Query Match 1.2%; Score 7; DB 12; Length 65;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGKV 433
|||||
Db 51 SRNLGKV 57

RESULT 52
ID Q68528 PRELIMINARY; PRT; 65 AA.
AC Q68528;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
```

```
DR EMBL; U45475; AAA86931.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Polyprotein.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 65
SQ SEQUENCE 65 AA; 7305 MW; A36A6E641472DE28 CRC64;

Query Match 1.2%; Score 7; DB 12; Length 65;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGKV 433
|||||
Db 51 SRNLGKV 57

RESULT 51
ID Q68527 PRELIMINARY; PRT; 65 AA.
AC Q68527;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Genome polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTx;
RA Mueller H.M., Selzer S., Heller A.E., Pfaff E., Theilmann L.;
RT "Recurrent hepatitis C virus infection after liver transplantation in
RT a patient negative for viral proteins: A possible role of immune
RT tolerance."
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -|- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; U45470; AAA86926.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Polyprotein.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 65
SQ SEQUENCE 65 AA; 7335 MW; A36A6E786472DE28 CRC64;

Query Match 1.2%; Score 7; DB 12; Length 65;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGKV 433
|||||
Db 51 SRNLGKV 57

RESULT 52
ID Q68528 PRELIMINARY; PRT; 65 AA.
AC Q68528;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
```

DR Pfam; PF01542; HCV_core; 1.
KW Polyprotein. 1 1
FT NON_TER 65 65
FT NON_TER 65 65
SQ SEQUENCE 65 AA; 7335 MW; A36A6E786472DE28 CRC64;
Query Match 1.2%; Score 7; DB 12; Length 65;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 427 SRNLGKV 433
Db 51 SRNLGKV 57
RESULT 46
Q68521 ID Q68521 PRELIMINARY; PRT; 65 AA.
AC Q68521
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Genome polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTX;
RA Mueller H.M., Selzer S., Heller A.E., Pfaff E., Theilmann L.;
RT "Recurrent hepatitis C virus infection after liver transplantation in
a patient negative for viral proteins: A possible role of immune
tolerance.";
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; U45464; AAA86920.1; -;
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
KW Polyprotein. 1 1
FT NON_TER 65 65
FT NON_TER 65 65
SQ SEQUENCE 65 AA; 7335 MW; A36A6E786472DE28 CRC64;
Query Match 1.2%; Score 7; DB 12; Length 65;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 427 SRNLGKV 433
Db 51 SRNLGKV 57
RESULT 47
Q68522 ID Q68522 PRELIMINARY; PRT; 65 AA.
AC Q68522
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Genome polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTX;
RA Mueller H.M., Selzer S., Heller A.E., Pfaff E., Theilmann L.;
RT "Recurrent hepatitis C virus infection after liver transplantation in
a patient negative for viral proteins: A possible role of immune
tolerance.";
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; U45464; AAA86920.1; -;
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
KW Polyprotein. 1 1
FT NON_TER 65 65
FT NON_TER 65 65
SQ SEQUENCE 65 AA; 7335 MW; A36A6E786472DE28 CRC64;
Query Match 1.2%; Score 7; DB 12; Length 65;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 427 SRNLGKV 433
Db 51 SRNLGKV 57
RESULT 48
Q68523 ID Q68523 PRELIMINARY; PRT; 65 AA.
AC Q68523
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Genome polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTX;
RA Mueller H.M., Selzer S., Heller A.E., Pfaff E., Theilmann L.;
RT "Recurrent hepatitis C virus infection after liver transplantation in
a patient negative for viral proteins: A possible role of immune
tolerance.";
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; U45464; AAA86922.1; -;
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
KW Polyprotein. 1 1
FT NON_TER 65 65
FT NON_TER 65 65
SQ SEQUENCE 65 AA; 7335 MW; A36A6E786472DE28 CRC64;
Query Match 1.2%; Score 7; DB 12; Length 65;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 427 SRNLGKV 433
Db 51 SRNLGKV 57

RP SEQUENCE FROM N.A.
RC STRAIN=HTX;
RA Mueller H.M., Selzer S., Heller A.E., Pfaff E., Theilmann L.;
RT "Recurrent hepatitis C virus infection after liver transplantation in
a patient negative for viral proteins: A possible role of immune
tolerance.";
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; U45465; AAA86921.1; -;
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
KW Polyprotein. 1 1
FT NON_TER 65 65
FT NON_TER 65 65
SQ SEQUENCE 65 AA; 7335 MW; A36A6E786472DE28 CRC64;
Query Match 1.2%; Score 7; DB 12; Length 65;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 427 SRNLGKV 433
Db 51 SRNLGKV 57
RESULT 49
Q68524 ID Q68524 PRELIMINARY; PRT; 65 AA.
AC Q68524
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Genome polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTX;
RA Mueller H.M., Selzer S., Heller A.E., Pfaff E., Theilmann L.;
RT "Recurrent hepatitis C virus infection after liver transplantation in
a patient negative for viral proteins: A possible role of immune
tolerance.";
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
DR EMBL; U45466; AAA86922.1; -;
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
KW Polyprotein. 1 1
FT NON_TER 65 65
FT NON_TER 65 65
SQ SEQUENCE 65 AA; 7335 MW; A36A6E786472DE28 CRC64;
Query Match 1.2%; Score 7; DB 12; Length 65;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 427 SRNLGKV 433
Db 51 SRNLGKV 57

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=226900;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22608415; PubMed=12721630;
 RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,
 RA Kapatral V., Shattacharya A., Reznik G., Mikhailova N., Lapidus A.,
 RA Chu L., Mazur M., Goltzman E., Larsen N., D'Souza M., Walunas T.,
 RA Grechkin Y., Pusch G., Haselkorn R., Fonstein M., Ehrlich S.D.,
 RA Overbeek R., Kyjpides N.;
 RT "Genome sequence of Bacillus cereus and comparative analysis with
 RT Bacillus anthracis";
 RL Nature 423:87-91(2003).
 DR EMBL; AF017004; AAP08984.1; --
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 63 AA; 7818 MW; A7A755FCE06B46B CRC64;

Query Match 1.2%; Score 7; DB 16; Length 63;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 515 Tlseker 521
 DB 29 Tlseker 35
 |||||

RESULT 43
 Q68530 PRELIMINARY; PRT; 65 AA.
 AC Q68530;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Genome polyprotein (Fragment).
 CC Lipoprotein envelope. The envelope consists of two proteins:
 CC Protein M and glycoprotein E. The nucleocapsid is a complex of
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HTX;
 RA Mueller H.M., Selzer S., Heller A.E., Pfaff E., Theilmann L.;
 RT "Recurrent hepatitis C virus infection after liver transplantation in
 RT a patient negative for viral proteins: A possible role of immune
 RT tolerance";
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
 CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
 CC PROTEIN C AND MRNA (BY SIMILARITY).
 DR EMBL; U45473; AAA86929.1; --
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR InterPro; IPR002522; HCV capsid.
 DR InterPro; IPR002521; HCV_core.
 DR Pfam; PF01543; HCV_capsid; 1.
 DR Pfam; PF01542; HCV_core; 1.
 KW Polyprotein.
 FT NON TER 1 1
 FT NON TER 65 65
 SQ SEQUENCE 65 AA; 7335 MW; A36A6E786472DE28 CRC64;

Query Match 1.2%; Score 7; DB 12; Length 65;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGV 433
 DB 51 SRNLGV 57
 |||||

RESULT 44
 Q68525 PRELIMINARY; PRT; 65 AA.
 AC Q68525;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Genome polyprotein (Fragment).
 CC Lipoprotein envelope. The envelope consists of two proteins:
 CC Protein M and glycoprotein E. The nucleocapsid is a complex of
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HTX;
 RA Mueller H.M., Selzer S., Heller A.E., Pfaff E., Theilmann L.;
 RT "Recurrent hepatitis C virus infection after liver transplantation in
 RT a patient negative for viral proteins: A possible role of immune
 RT tolerance";
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
 CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
 CC PROTEIN C AND MRNA (BY SIMILARITY).
 DR EMBL; U45467; AAA86923.1; --
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR InterPro; IPR002522; HCV capsid.
 DR InterPro; IPR002521; HCV_core.
 DR Pfam; PF01543; HCV_capsid; 1.
 DR Pfam; PF01542; HCV_core; 1.
 KW Polyprotein.
 FT NON TER 1 1
 FT NON TER 65 65
 SQ SEQUENCE 65 AA; 7335 MW; A36A6E786472DE28 CRC64;

Q68525 PRELIMINARY; PRT; 65 AA.
 AC Q68525;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Genome polyprotein (Fragment).
 CC Lipoprotein envelope. The envelope consists of two proteins:
 CC Protein M and glycoprotein E. The nucleocapsid is a complex of
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HTX;
 RA Mueller H.M., Selzer S., Heller A.E., Pfaff E., Theilmann L.;
 RT "Recurrent hepatitis C virus infection after liver transplantation in
 RT a patient negative for viral proteins: A possible role of immune
 RT tolerance";
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
 CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
 CC PROTEIN C AND MRNA (BY SIMILARITY).
 DR EMBL; U45468; AAA86924.1; --
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR InterPro; IPR002522; HCV capsid.
 DR InterPro; IPR002521; HCV_core.
 DR Pfam; PF01543; HCV_capsid; 1.
 DR Pfam; PF01542; HCV_core; 1.
 KW Polyprotein.
 FT NON TER 1 1
 FT NON TER 65 65
 SQ SEQUENCE 65 AA; 7335 MW; A36A6E786472DE28 CRC64;

Query Match 1.2%; Score 7; DB 12; Length 65;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 SRNLGV 433
 DB 51 SRNLGV 57
 |||||

RESULT 45
 Q68524 PRELIMINARY; PRT; 65 AA.
 AC Q68524;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Genome polyprotein (Fragment).
 CC Lipoprotein envelope. The envelope consists of two proteins:
 CC Protein M and glycoprotein E. The nucleocapsid is a complex of
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HTX;
 RA Mueller H.M., Selzer S., Heller A.E., Pfaff E., Theilmann L.;
 RT "Recurrent hepatitis C virus infection after liver transplantation in
 RT a patient negative for viral proteins: A possible role of immune
 RT tolerance";
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
 CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
 CC PROTEIN C AND MRNA (BY SIMILARITY).
 DR EMBL; U45467; AAA86923.1; --
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR InterPro; IPR002522; HCV capsid.
 DR InterPro; IPR002521; HCV_core.
 DR Pfam; PF01543; HCV_capsid; 1.
 DR Pfam; PF01542; HCV_core; 1.


```

RESULT 39
Q9LOK8 PRELIMINARY; PRT; 1090 AA.
ID Q9LOK8
AC Q9LOK8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE F5D14.27 protein.
DE F5D14.27.
GN Arabidopsis thaliana (Mouse-ear cress).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Liu S.X., Chan A., Yu G., Lee J.M., Lenz C., Pham P., Sakano H.,
RA Toriumi M., Vysotskaya V.S., Chin C., Chioi J., Choi E., Chung M.,
RA Gonzalez A., Hong B., Liu A., Vaysberg M., Altati H., Brooks S.,
RA Buehler E., Chao Q., Conn L., Conway A.B., Hansen N.F.,
RA Johnson-Hopson C., Khan S., Kim C., Lam B., Miranda M., Nguyen M.,
RA Palm C.J., Shinn P., Southwick A., Davis R.W., Ecker J.R.,
RA Pederspiel N.A., Theologis A.;
RA "The sequence of BAC F5D14 from Arabidopsis thaliana chromosome 1."
RT Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Theologis A.;
RL EMBL; AC007767; AAF81347.1; -.
RL PIR; C86450; C86450.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0008026; P:ATP dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002464; DEAD_box.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR007502; Helicase_dom.
DR Pfam; PF04408; HX2; 1.
DR Pfam; PF00271; helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC2; 1.
DR PROSITE; PS00690; DEAD_ATP_HELICASE; 1.
DR ATP-binding; Helicase; Hydrolase.
KW SEQUENCE 1090 AA; 124147 MW; FBA556887DD06EE CRC64;
SQ
Query Match 1.4%; Score 8; DB 10; Length 1090;
Best Local Similarity 100.0%; Pred.No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps
QY 181 KLDELRLDE 188
DB 241 KLDELRLDE 248
|||||||
|||||||

RESULT 40
Q7UL51 PRELIMINARY; PRT; 36 AA.
ID Q7UL51
AC Q7UL51;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
DE R59736.
GN Rhodospirillum rubrum.
OS Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OC NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.

```

RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAOI;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.D., Huynh W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Golltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of *Pseudomonas aeruginosa* PAOI, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
RL EMBL; AE004466; AAG03674.1; -.
DR PIR; E83610; E83610.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR001633; EAL.
DR InterPro; IPR001600; GDEF.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000700; PAS-associ C.
DR InterPro; IPR000114; PAS_domain.
DR Pfam; PF00563; EAL; 1.
DR Pfam; PF00990; GDEF; 1.
DR Pfam; PF00785; PAC; 2.
DR Pfam; PF00989; PAS; 2.
DR SMART; SM00267; DUF1; 1.
DR SMART; SM00052; DUF2; 1.
DR SMART; SM00086; PAC; 2.
DR SMART; SM00091; PAS; 2.
DR TIGFams; TIGR00254; GDEF; 1.
DR TIGFams; TIGR00229; sensory_box; 2.
DR PROSITE; PS50883; EAL; 1.
DR PROSITE; PS50887; GDEF; 1.
DR PROSITE; PS50113; PAC; 2.
DR PROSITE; PS50112; PAS; 2.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 760 AA; 86380 MW; 049AFD1D6E222E05 CRC64;

Query Match 1.4%; Score 8; DB 16; Length 760;
Best Local Similarity 100.0%; Pred.No. 85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 AVARLSQR 222
DB 425 AVARLSQR 432
|||||

RESULT 37
QYV975 PRELIMINARY; PRT; 948 AA.
AC QYV975;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Preprotein translocase SecA subunit.
GN SECA OR PMT0083.
OS *Prochlorococcus marinus* (strain MIT 9313).
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC *Prochlorococcus*.
OX NCBI_TaxID=74547;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22826598; PubMed=12917642;
RA Roca P.G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinser E.R., Chisholm S.W.;
RT "Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic
RT niche differentiation.";
RL Nature 424:1042-1047(2003).
RL EMBL; BX572095; CAE20258.1; -.
KW Complete proteome.

SQ SEQUENCE 948 AA; 107136 MW; ABD79A189DEDEDE0 CRC64;

Query Match 1.4%; Score 8; DB 16; Length 948;
Best Local Similarity 100.0%; Pred.No. 1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 LPRVLRPE 119
DB 529 LPRVLRPE 536
|||||

RESULT 38
QYVY00 PRELIMINARY; PRT; 1044 AA.
ID QYVY00
AC QYVY00;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative RNA helicase.
GN ATIG32490.
OS *Arabidopsis thaliana* (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Ban J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA Davis R.W., Ecker J.R., Theologis A.;
RT "Full length cDNA of gene Atig32490 (GI:15223221).";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Deng J.M., Huan V.W.,
RA Lee J.M., Quach H.L., Tang C., Toriumi M., Wallender E.K., Wong C.,
RA Wu H.C., Yu G., Yuan S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Lam B., Lin J., Miranda M., Narusaka M., Nguyen M.,
RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "Arabidopsis Open Reading Frame (ORF) Clones";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY074318; AAL67014.1; -.
DR EMBL; AY133872; AAM91806.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002464; DEAH_box.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR007502; Helicase_dom.
DR Pfam; PF04408; HA2; 1.
DR Pfam; PF00271; Helicase_C_1.
DR PROSITE; PS00690; DEAH_ATP_HELICASE; 1.
KW ATP-binding; Helicase; Hydrolase.
SQ SEQUENCE 1044 AA; 118831 MW; 187B02E796AF0E18 CRC64;

Query Match 1.4%; Score 8; DB 10; Length 1044;
Best Local Similarity 100.0%; Pred.No. 1.1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 181 KLDELRLDE 188
DB 237 KLDELRLDE 244
|||||

OX NCBI_TaxID=7227;
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutcliffe G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fodor C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AF003579; AAF51091.1; -
DR FlyBase; FBGN0031548; CG8852.
DR InterPro; IPR001611; LRR.
DR Pfam; PF00560; LRR; 3
SQ SEQUENCE 663 AA; 74350 MW; 379D3B185854C436 CRC64;

Query Match 1.4%; Score 8; DB 5; Length 663;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 PKLDELRD 187
DB 654 PKLDELRD 661

RESULT 34
Q8SZ97 PRELIMINARY; PRT; 663 AA.
AC Q8SZ97;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Re11035p.
GN CG8852.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Friese E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nurco J., Pacle J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY071028; AAL48650.1; -
DR FlyBase; FBGN0031548; CG8852.
DR InterPro; IPR001611; LRR.
DR Pfam; PF00560; LRR; 3
SQ SEQUENCE 663 AA; 74369 MW; 3BECFC01DB5DB4A1 CRC64;

Query Match 1.4%; Score 8; DB 5; Length 663;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 PKLDELRD 187
DB 654 PKLDELRD 661

RESULT 35
Q9JG63 PRELIMINARY; PRT; 752 AA.
AC Q9JG63;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Brain cDNA, clone MNCB-4173.
GN D8ERTD594E OR C85457.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Osada N., Kusuda J., Tanuma R., Ito A., Hirata M., Sugano S.,
RA Hashimoto K.;
RL "Isolation of full-length cDNA clones from mouse brain cDNA library.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB045323; BAA97983.1; -
DR MGD; MGI:1261872; D8ERTD594E.
DR InterPro; IPR000008; C2.
DR Pfam; PF00168; C2; 1.
SQ SEQUENCE 752 AA; 82720 MW; CF27BE23479AADBD CRC64;

Query Match 1.4%; Score 8; DB 11; Length 752;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 576 VAASQAAL 583
DB 336 VAASQAAL 343

RESULT 36
Q916K5 PRELIMINARY; PRT; 760 AA.
AC Q916K5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein PA0285.
GN PA0285.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]

RA Salzberg S.L., White O., Fraser C.M.;
 RT "Oryza sativa chromosome 3 BAC OSJNB0048D20 genomic sequence."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC084766; AAL82518.1; -;
 DR Gramineae; Q8S7F6; -;
 KW Hypothetical protein.
 SQ SEQUENCE 595 AA; 67436 MW; A9394FD97CA8BD08 CRC64;

Query Match 1.4%; Score 8; DB 10; Length 595;
 Best Local Similarity 100.0%; Pred.No. 69;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 DKSLHTLF 70
 |||||
 Db 193 DKSLHTLF 200

RESULT 30
 Q8S7F6 PRELIMINARY; PRT; 602 AA.
 AC Q8S7F6;
 DT 01-JUN-2002 (TREMELrel. 21, Created)
 DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN OSJNB0048D20.3.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
 RA Ganeberg K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
 RA Tsirkin T., Riggs P., Heiao J., Zismann V., Blunt S., Pai G.,
 RA Vanaken S.E., Utterback T.R., Feldblyum T.V., Kalb E., Quackenbush J.,
 RA Salzberg S.L., White O., Fraser C.M.;
 RT "Oryza sativa chromosome 3 BAC OSJNB0048D20 genomic sequence."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC084766; AAL82526.1; -;
 DR Gramineae; Q8S7F6; -;
 KW Hypothetical protein.
 SQ SEQUENCE 602 AA; 68561 MW; 4CDFD59ECFF99F00 CRC64;

Query Match 1.4%; Score 8; DB 10; Length 602;
 Best Local Similarity 100.0%; Pred.No. 69;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 DKSLHTLF 70
 |||||
 Db 148 DKSLHTLF 155

RESULT 31
 Q8MJU5 PRELIMINARY; PRT; 609 AA.
 AC Q8MJU5;
 DT 01-OCT-2002 (TREMELrel. 22, Created)
 DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Alpha-fetoprotein.
 GN AFP.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Furuichi M., Neo S., Hisasue M., Tsuchiya R., Watanabe M.,
 RA Hashizaki K., Hisamatsu S., Yamada T.;
 RT "Canine alpha-fetoprotein cDNA.";

RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB089789; BAC07513.1; -;
 DR GO; GO:0005615; C:extracellular space; IEA.
 DR GO; GO:0005386; F:carrier activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000264; Serum albumin.
 DR Pfam; PF00273; transport_prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR PRODOM; PD002486; Serum albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 2.
 SQ SEQUENCE 609 AA; 68782 MW; BE4B8250CSAF2AF0 CRC64;

Query Match 1.4%; Score 8; DB 6; Length 609;
 Best Local Similarity 100.0%; Pred.No. 70;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 YEIARRHP 147
 |||||
 Db 164 YEIARRHP 171

RESULT 32
 Q8MJ76 PRELIMINARY; PRT; 610 AA.
 ID Q8MJ76
 AC Q8MJ76;
 DT 01-OCT-2002 (TREMELrel. 22, Created)
 DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Alpha-fetoprotein.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Kim J.G., Nonnenan D., Vallet J.L., Christensen R.K.;
 RA "Mapping of the porcine alpha-fetoprotein (AFP) gene to SSC8."
 RT Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF517770; AAM66710.1; -;
 DR GO; GO:0005615; C:extracellular space; IEA.
 DR GO; GO:0005386; F:carrier activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000264; Serum albumin.
 DR Pfam; PF00273; transport_prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR PRODOM; PD002486; Serum albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 2.
 SQ SEQUENCE 610 AA; 68624 MW; C985BEAD44963D5E CRC64;

Query Match 1.4%; Score 8; DB 6; Length 610;
 Best Local Similarity 100.0%; Pred.No. 70;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 YEIARRHP 147
 |||||
 Db 164 YEIARRHP 171

RESULT 33
 Q9VQS3 PRELIMINARY; PRT; 663 AA.
 ID Q9VQS3
 AC Q9VQS3;
 DT 01-MAY-2000 (TREMELrel. 13, Created)
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE CG8852 protein.
 GN CG8852.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.

AC QXBL3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PEP-protein phosphotransferase system enzyme I.
GN PFSI OR Z3682 OR ECS3288.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glaesner J.D.,
Rose D.J., Mayhew G.F., Evans F.S., Gregor J., Kirkpatrick H.A.,
Rosa J.G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
Rafaeli A., Davis N.W., Lim A., Dimalanta E.T., Potamocitis K.,
Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RL Nature 409:529-533 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Onishi M., Kurokawa K., Iehi K., Yokoyama K.,
Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
Kuhara S., Shiba T., Hattori M., Shinagawa H.,
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12";
RL DNA Res. 8:11-22 (2001).
DR ENBL; AP002561; BAB36711.1; -;
DR PIR; C85884; C85884.
DR PIR; H91039; H91039.
DR GO; GO:0016772; F:transferase activity, transferring phosphor. .; IEA.
DR GO; GO:0016310; P:phosphorylation; IEA.
DR InterPro; IPR008731; PEP-utilizers_N.
DR InterPro; IPR008279; PEP mobile.
DR InterPro; IPR006318; PEP_P trans.
DR InterPro; IPR000121; PEP-utilizers.
DR Pfam; PF005524; PEP-utilizers_N; 1.
DR Pfam; PF00391; PEP-utilizers; 1.
DR Pfam; PF02896; PEP-utilizers_C; 1.
DR PRINTS; PR01736; PHPTRNFRASE.
DR ProDom; PD000940; PEP-utilizers; 1.
DR TIGRFAMs; TIGR01417; PTS_I_fam; 1.
DR PROSITE; PS00742; PEP-ENZYMES_2; 1.
DR PROSITE; PS00370; PEP-ENZYMES_PHOS_SITE; 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 575 AA; 63533 MW; B9C5344D6AE827B2 CRC64;
Query Match 1.4%; Score 8; DB 16; Length 575;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 184 ELRDEGKA 191
DB 412 ELRDEGKA 419
RESULT 28
Q83K79
ID Q83K79 PRELIMINARY; PRT; 575 AA.
AC Q83K79;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE PEP-protein phosphotransferase system enzyme I.
GN PFSI OR SF2471 OR S2617.
OS Shigella flexneri.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=2227406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
through comparison with genomes of Escherichia coli K12 and O157";
RL Nucleic Acids Res. 30:4432-4441 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=1270415; Venkatesan M.M., Deng W.,
Wei J., Goldberg M.B., Burland V., Plunkett G. III, Rose D.J., Darling A.,
Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
flexneri serotype 2a strain 2457T";
RL Infect. Immun. 71:2775-2786 (2003).
DR ENBL; AE015262; AAN43978.1; -;
DR GO; GO:0016772; F:transferase activity, transferring phosphor. .; IEA.
DR GO; GO:0016310; P:phosphorylation; IEA.
DR InterPro; IPR008279; PEP mobile.
DR InterPro; IPR006318; PEP_P trans.
DR InterPro; IPR000121; PEP-utilizers.
DR Pfam; PF00391; PEP-utilizers; 1.
DR Pfam; PF02896; PEP-utilizers_C; 1.
DR PRINTS; PR01736; PHPTRNFRASE.
DR ProDom; PD000940; PEP-utilizers; 1.
DR TIGRFAMs; TIGR01417; PTS_I_fam; 1.
DR PROSITE; PS00742; PEP-ENZYMES_2; 1.
DR PROSITE; PS00370; PEP-ENZYMES_PHOS_SITE; 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 575 AA; 63533 MW; B9C5344D6AE827B2 CRC64;
Query Match 1.4%; Score 8; DB 16; Length 575;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 184 ELRDEGKA 191
DB 412 ELRDEGKA 419
RESULT 29
Q8S7F9
ID Q8S7F9 PRELIMINARY; PRT; 595 AA.
AC Q8S7F9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OSUNB0048D20.2
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
Gansberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
Taitrin T., Riggs F., Hsiao J., Zismann V., Blunt S., Pal G.,
Vanaken S.E., Utterback T.R., Feldblyum T.V., Kalb E., Quackenbush J.,

```

RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882 (2001).
DR EMBL; AL591792; CAC47620.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000792; HTH_LuxR.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 509 AA; 55401 MW; 2FA8F0B1EEOA9362 CRC64;

Query Match 1.4%; Score 8; DB 16; Length 509;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 333 EYARRHPD 340
Db 323 EYARRHPD 330
|||||

RESULT 26
ID Q06729 PRELIMINARY; PRT; 538 AA.
AC Q06729;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE MAT locus genes BUD5, MAT-ALPHA1, MAT-ALPHA2, YCR724 and YCR725.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=AB972;
RX MEDLINE=92160397; PubMed=1789011;
RA Jacquet M., Buhler J.M., Iborra F., Francinques-Gaillard M.C.,
RA Soustelle C.;
RT "The MAT locus revisited within a 9.8 kb fragment of chromosome III
RL containing BUD5 and two new open reading frames.";
RL Yeast 7:881-888 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB972;
RA Jacquet M.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; X63853; CAA45334.1; -.
DR PIR; S19450; BWBYD5.
DR GO; GO:0005935; C:bud neck; IEA.
DR GO; GO:0001313; C:incipient bud site; IEA.
DR InterPro; IPR000651; RasGEF.
DR InterPro; IPR001895; RasGRP_CDC25.
DR InterPro; IPR008937; Ras_GEF.
DR Pfam; PF00617; RasGEF; 1.
DR SMART; SM00147; RasGEF; 1.
DR SMART; SM00229; RasGEF; 1.
DR PROSITE; PS00720; RasGEF; 1.
DR PROSITE; PS00009; RasGEF_CAT; 1.
DR PROSITE; PS00212; RasGEF_NTER; 1.
DR PROSITE; PS00212; RasGEF_NTER; 1.
SQ SEQUENCE 538 AA; 62869 MW; F3D0C613D0C0BE31 CRC64;

Query Match 1.4%; Score 8; DB 3; Length 538;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AFAQYLQ 33
Db 485 AFAQYLQ 492
|||||

RESULT 27
Q8XBL3 PRELIMINARY; PRT; 575 AA.
ID Q8XBL3

```

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin C.M., Venter J.C.;
RA "The genome sequence of *Drosophila melanogaster*."
RT Science 287:2185-2195(2000).
RL [2]
RN
RP
SEQUENCE FROM N.A.
RA Amanatides P.G., Brandon R.C., Rogers Y., An H., Baldwin D.,
RA Banazon J., Besson K.Y., Bueam D.A., Carlson J.W., Center A.,
RA Champ M., Davenport L.B., Dietz S.M., Dodson K., Dorsett V.,
RA Doup L.E., Doyle C., Dreanek D., Farfan D., Ferreira S., Frise E.,
RA Galle R.F., Garg N.S., George R.A., Gonzalez M., Houck J.,
RA Hoskins R.A., Hostin D., Howland T.J., Ibegwam C., Jalali M.,
RA Kruse D., Li P., Mattei B., Mohrrefti A., McIntosh T.C., Moy M.,
RA Murphy B., Nelson C., Nelson K.A., Nunco J., Pacleb J., Paragas V.,
RA Park S., Patel S., Pfeiffer B., Phouanavong S., Pittman G.S.,
RA Puri V., Richards S., Scheeler P., Stapleton M., Strong R.,
RA Svirskas R., Tector C., Tyler D., Williams S.M., Zaveri J.S.,
RA Smith H.O., Venter J.C.; Rubin G.M.;
RA "Sequencing of *Drosophila melanogaster* genome."
RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RL [3]
RN
RP
SEQUENCE FROM N.A.
RA Miera S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaniker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C.M., Berman B.P., Carlson J.W., Celniker S.E.,
RA Clamp M.B., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,
RA Harris N.L., Kronmiller B., Marshall B., Millburn G.H., Richter J.,
RA Russo S., Seale S.M.J., Smith E., Shu S., Smutniak F.,
RA Whitfield E.J., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Mungall C.J., Lewis S.E.;
RA "Annotation of *Drosophila melanogaster* genome."
RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RL
DR
EMBL; AB03458; AAM71129.1; -.
DR
FLYBase; FBgn0050274; CG30274.
DR
GO; GO:0016301; F:kinase activity; IEA.
DR
GO; GO:0016772; F:transferase activity, transferring phosphor. .; IEA.
DR
InterPro; IPR000749; ATP-gua Ptrans.
DR
Pfam; PF00217; ATP-gua Ptrans; 1.
DR
SEQUENCE 468 AA; 53055 MW; 4D877712E24A0BD9 CRC64;
SQ
Query Match 1.4%; Score 8; DB 5; Length 468;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 423 LVEVSRNL 430
DB 196 LVEVSRNL 203

RESULT 22
Q971B5 PRELIMINARY; PRT; 483 AA.
AC Q971B5
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE Putative prolyl-tRNA synthetase.
GN ST1440.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP
SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX MEDLINE=21456156; PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Anai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,

RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermocacidophilic
RT Crenarchaeon, *Sulfolobus tokodaii* strain7."
RL DNA Res. 8:123-140(2001).
DR EMBL; AP000986; BAB66508.1; -.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004827; F:proline-tRNA ligase activity; IEA.
DR GO; GO:0006433; F:prolyl-tRNA aminoacylation; IEA.
DR InterPro; IPR004154; HGTP anticodon.
DR InterPro; IPR004499; ProS fam I.
DR InterPro; IPR002314; tRNA-synt 2b.
DR InterPro; IPR002316; tRNA-synt-pro.
DR InterPro; IPR006195; tRNA ligase II.
DR Pfam; PF03129; HGTP anticodon; 1.
DR Pfam; PF00587; tRNA-synt 2b; 1.
DR PRINTS; PR01046; TRNASYNTHPRO.
DR PROSITE; PS00408; ProS fam I; 1.
DR PROSITE; PS00862; AA tRNA LIGASE II; 1.
KW Aminoacyl-tRNA synthetase; Hypothetical protein; Complete proteome.
SQ SEQUENCE 483 AA; 55921 MW; 8D05FEB74364B07A CRC64;
Query Match 1.4%; Score 8; DB 17; Length 483;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 346 LURLAKTY 353
DB 476 LURLAKTY 483

RESULT 23
Q8XL35 PRELIMINARY; PRT; 490 AA.
ID Q8XL35
AC Q8XL35
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
DE Hypothetical protein CPE1207.
GN CPE1207.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP
SEQUENCE FROM N.A.
RC STRAIN=13 / Type A;
RX MEDLINE=21664373; PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Saiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of *Clostridium perfringens*, an anaerobic
RT flesh-eater."
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL; AP003189; BAB80913.1; -.
DR InterPro; IPR007229; NAPTase.
DR InterPro; IPR006405; Nic Ptrans-like.
DR Pfam; PF04095; NAPTase; 1.
DR TIGRfams; TIGR01513; NAPTase; 1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 490 AA; 55929 MW; 58D5B84DEF039F62 CRC64;
Query Match 1.4%; Score 8; DB 16; Length 490;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 210 AFKAWAVA 217
DB 220 AFKAWAVA 227

RESULT 24

RC STRAIN=Massachusetts / B88;
RX MEDLINE=22457253; PubMed=12552129;
RA Brueggemann H., Baumer S., Fricke W.F., Wier A., Liesegang H.,
RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
RA Gottschalk G.;
RT "The genome sequence of Clostridium tetani, the causative agent of
RT tetanus disease.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
DR EMBL; AE015945; AA037019.1; -;
GO GO:0016829; F:lyase activity; IEA.
KW Lyase; Complete proteome.
SQ SEQUENCE 295 AA; 32503 MW; 9BBFF7B30C5B211 CRC64;
Query Match 1.4%; Score 8; DB 16; Length 295;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 537 PKATKEQL 544
DB 98 PKATKEQL 105
RESULT 19
Q8ROVS PRELIMINARY; PRT; 398 AA.
AC Q8ROVS; (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE Hypothetical protein.
GN C230043N17RIK OR A1265623.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC026393; AAH26393.1; -;
DR MGD; MGI:2142489; C230043N17RIK.
DR InterPro; IPR000898; IDO fam.
DR Pfam; PF01231; IDO; 1.
KW Hypothetical protein.
SQ SEQUENCE 398 AA; 44439 MW; 7871CCB86FE24D1A CRC64;
Query Match 1.4%; Score 8; DB 11; Length 398;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 424 VEVSRLNG 431
DB 109 VEVSRLNG 116
RESULT 20
O86640 PRELIMINARY; PRT; 422 AA.
ID O86640; (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein SC05720.
GN SC05720 OR SC3C3.06C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=A3(2) / M145;
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;

RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL399124; CAJ20255.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 422 AA; 45397 MW; 4ASB6B51953F71A3 CRC64;
Query Match 1.4%; Score 8; DB 16; Length 422;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 ALVLIAPA 28
DB 67 ALVLIAPA 74
RESULT 21
Q8MMD7 PRELIMINARY; PRT; 468 AA.
ID Q8MMD7; (TREMBlrel. 22, Created)
AC Q8MMD7; (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE Hypothetical protein.
GN CG30274-PA.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Beckley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Gbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burkis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RP SEQUENCE FROM N.A.
 RX MEDLINE=20493557; PubMed=10924517;
 RA Lu R., Au W.-C., Yew W.-S., Hageman N., Pitha P.M.;
 RT "Regulation of the promoter activity of interferon regulatory factor-7
 gene. ACTIVATION BY INTERFERON AND SILENCING BY HYPERMETHYLATION.";
 RL J. Biol. Chem. 275:31805-31812(2000).
 DR EMBL: AF277159; AAG30003.1; -
 DR HSSP: P23306; 21RF
 DR GO: GO:0005634; C:nucleus; IEA.
 DR GO: GO:0003700; F:transcription factor activity; IEA.
 DR GO: GO:0006355; F:regulation of transcription, DNA-dependent; IEA.
 DR InterPro: IPR001346; IRF.
 DR Pfam: PF00605; IRF; 1.
 DR PRINTS: PR00267; INTERREGPECT.
 DR ProDom: PD002355; IRF; 1.
 DR SMART: SM00348; IRF; 1.
 DR PROSITE: PS00801; IRF; 1.
 FT NON_TER 128 128
 SQ SEQUENCE 128 AA; 14363 MW; 09B5B6F3CC11BB15 CRC64;

Query Match 1.4%; Score 8; DB 4; Length 128;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 FKAWAVAR 218
 DB 73 FKAWAVAR 80

RESULT 16
 Q9PHJ3 PRELIMINARY; PRT; 273 AA.
 ID Q9PHJ3
 AC Q9PHJ3
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Confugal transfer protein.
 GN XFA0012.
 OS Xylella fastidiosa.
 OG Plasmid pXf51.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xylella.
 OX NCBI_TaxID=2317;
 [1]
 RN RP
 RP SEQUENCE FROM N.A.
 RC STRAIN=9a5C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvares A.J.G., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.J.S., Frohme M., Furlan L.R.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hobeisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.B., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.N., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva P.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Silveira W.J., de Souza A.A.,
 RA de Souza A.P., Terezi M.F., Truffi D., Teai S.M., Tshako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;

RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
 RL Nature 406:151-159(2000).
 DR EMBL: AE003851; AAF85581.1; -
 DR PIR: D82862; D82862.
 DR InterPro: IPR007430; VirB8.
 DR Pfam: PF04335; VirB8; 1.
 KW Complete proteome.
 SQ SEQUENCE 273 AA; 30593 MW; B6992B5F8D389A0A CRC64;

Query Match 1.4%; Score 8; DB 16; Length 273;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 408 LVRYTKKV 415
 DB 170 LVRYTKKV 177

RESULT 17
 P74538 PRELIMINARY; PRT; 289 AA.
 ID P74538
 AC P74538
 DT 01-FEB-1997 (TRENBLrel. 02, Created)
 DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Hypothetical protein slr1429.
 GN SLR1429.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 [1]
 RN RP
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=9905231;
 RA Kaneko T., Sato S., Kocani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hikosawa M., Sugita M., Shimada K., Shimada K.,
 RA Hosouchi T., Hatakeyama A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
 RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 DR EMBL: D90916; BAA18644.1; -
 DR PIR: S76732; S76732.
 DR InterPro: IPR007527; Znf_SWIM.
 DR Pfam: PF04434; SWIM; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 289 AA; 32818 MW; 0465E15B15F6F2F5 CRC64;

Query Match 1.4%; Score 8; DB 16; Length 289;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 577 AASQAALG 584
 DB 275 AASQAALG 282

RESULT 18
 Q89056 PRELIMINARY; PRT; 295 AA.
 ID Q89056
 AC Q89056;
 DT 01-JUN-2003 (TRENBLrel. 24, Created)
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
 DE Citrate lyase beta chain.
 GN CTC02560.
 OS Clostridium tetani.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1513;
 [1]
 RN RP
 RP SEQUENCE FROM N.A.

OX NCBI_TaxID=10118;
RN SEQUENCE.
RX MEDLINE=96087830; PubMed=8587647;
RA Onozuka M., Inai S., Isobe T., Yen C.T., Watanabe K.;
RT "Purification and characterization of a novel 70-kDa brain protein
RT associated with seizure activities.";
RL Neurochem. Res. 20:901-905(1995).
DR HSP, 202768; 1E7H.
SQ SEQUENCE 20 AA; 2381 MW; 53423C0F9F70F7D CRC64;

Query Match 1.5%; Score 9; DB 11; Length 20;
Best Local Similarity 100.0%; Pred.No. 0.32;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AHRFKDLGE 16
Db 8 AHRFKDLGE 16

RESULT 12
Q9TQZ6 PRELIMINARY; PRT; 20 AA.
AC Q9TQZ6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE Albumin (Fragment).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
ON NCBI_TaxID=9541;
RN [1]
RP SEQUENCE.
RX MEDLINE=96273610; PubMed=8690030;
RA Nicolas M.G., Fujiki K., Murayama K., Suzuki M.T., Mineki R.,
RA Hayakawa M., Yoshikawa Y., Cho F., Kanai A.;
RT "Studies on the mechanism of early onset macular degeneration in
RT cynomolgus (Macaca fascicularis) monkeys. I. Abnormal concentrations of
RT two proteins in the retina."
RL Exp. Eye Res. 62:211-219(1996).
SQ SEQUENCE 20 AA; 2411 MW; 5F1A6AEB5918F777 CRC64;

Query Match 1.4%; Score 8; DB 6; Length 20;
Best Local Similarity 100.0%; Pred.No. 3.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HKSEVAHR 10
Db 13 HKSEVAHR 20

RESULT 13
Q95MC2 PRELIMINARY; PRT; 44 AA.
AC Q95MC2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Albumin (fragment).
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
ON NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2134992; PubMed=11421942;
RA Shubitowski D.M., Venna P.J., Douglass C.L., Zhou R.-X., Ewart S.L.;
RT "Polymorphism identification within 50 equine gene-specific sequence
RT tagged sites."
RL Anim. Genet. 32:78-78(2001).
DR ENBL; AY008769; AAG40944.1; -.

DR GO:0005615; C:extracellular space; IEA.
DR GO:0005386; F:carrier activity; IEA.
DR GO:0006810; P:transport; IEA.
DR InterPro: IPR000264; Serum albumin.
DR Pfam: PF00273; transport_prot; 1.
FT NON_TER 1
FT NON_TER 44
SQ SEQUENCE 44 AA; 5160 MW; 50C00290E137D646 CRC64;

Query Match 1.4%; Score 8; DB 6; Length 44;
Best Local Similarity 100.0%; Pred.No. 6.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 PERNECFL 103
Db 1 PERNECFL 8

RESULT 14
Q8QML2 PRELIMINARY; PRT; 115 AA.
AC Q8QML2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Genome polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
ON NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=928556a;
RA Theamboonlers A., Bedi K., Scottaj N., Sriponthong M., Poovorawan Y.;
RT "Hepatitis C virus genotyping by RFLP and direct sequencing in
RT Thailand.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND RNA (BY SIMILARITY).
DR EMBL; AY089756; AM09916.1; -.
DR GO:0019028; C:viral capsid; IEA.
DR GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002522; HCV capsid.
DR Pfam: PF01543; HCV_capsid; 1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 115
SQ SEQUENCE 115 AA; 12867 MW; BC0F619B38A21379 CRC64;

Query Match 1.4%; Score 8; DB 12; Length 115;
Best Local Similarity 100.0%; Pred.No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 426 VSRNLGV 433
Db 104 VSRNLGV 111

RESULT 15
Q9HB64 PRELIMINARY; PRT; 128 AA.
AC Q9HB64;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Interferon regulatory factor-7H (fragment).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RN [1]

```
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK050644; BAC34360.1; -.
DR MGD; MGI:87991; Albl.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
FT NON_TER 1
SQ SEQUENCE 576 AA; 65002 MW; F85733E9AE37F04 CRC64;

Query Match 3.1%; Score 18; DB 11; Length 576;
Best Local Similarity 100.0%; Pred. No. 3.5e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 KQEPERNECFLOHKDDNP 110
Db 85 KQEPERNECFLOHKDDNP 102

RESULT 8
Q8C7H3 PRELIMINARY; PRT; 608 AA.
ID Q8C7H3
AC Q8C7H3
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Albumin 1.
GN ALB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK050248; BAC34145.1; -.
DR MGD; MGI:87991; Albl.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
SQ SEQUENCE 608 AA; 68722 MW; 292F600BED3A61B4 CRC64;

Query Match 3.1%; Score 18; DB 11; Length 608;
Best Local Similarity 100.0%; Pred. No. 3.5e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 KQEPERNECFLOHKDDNP 110
Db 117 KQEPERNECFLOHKDDNP 134

RESULT 9
Q9TS75 PRELIMINARY; PRT; 61 AA.
ID Q9TS75
```

```
AC Q9TS75;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE SG-BSPOT Group B protein serum albumin homolog (Fragments).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE.
RX MEDLINE=95007849; PubMed=7923441;
RA Miller M.J.; Parmelee D.C.; Benjamin T.; Sechi S.; Dooley K.L.;
RA Kadiubar F.F.;
RT "Plasma proteins as early biomarkers of exposure to carcinogenic
RT aromatic amines."
RL Chem.-Biol. Interact. 93:221-234 (1994).
FT NON_TER 1
FT NON_CONS 28 29
FT NON_CONS 45 46
FT NON_TER 61
SQ SEQUENCE 61 AA; 7009 MW; F5B74797822D725 CRC64;

Query Match 1.7%; Score 10; DB 6; Length 61;
Best Local Similarity 100.0%; Pred. No. 0.081;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 503 NAEFTTFHAD 512
Db 48 NAEFTTFHAD 57

RESULT 10
Q9R4X7 PRELIMINARY; PRT; 20 AA.
ID Q9R4X7
AC Q9R4X7
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE 66 kDa SEROREACTIVE protein/serum albumin homolog (Fragment).
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE.
RX MEDLINE=94343500; PubMed=8064836;
RA Deshpande R.G.; Khan M.B.; Bhat D.A.; Navalkar R.G.;
RA "Purification and partial characterisation of a novel 66-kDa
RT seroreactive protein of Mycobacterium tuberculosis H37Rv."
RL J. Med. Microbiol. 41:173-178 (1994).
SQ SEQUENCE 20 AA; 2393 MW; 534A232072870F7D CRC64;

Query Match 1.5%; Score 9; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AHRFKDLGE 16
Db 8 AHRFKDLGE 16

RESULT 11
Q9QUX8 PRELIMINARY; PRT; 20 AA.
ID Q9QUX8
AC Q9QUX8
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE 70 kDa seizure activity-linked albumin-like glycoprotein (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
```

DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PRO00802; Serum_albumin; 1.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 608 AA; 68225 MW; E5EAB2B8E1C66E54 CRC64;

Query Match 4.6%; Score 27; DB 5; Length 608;
Best Local Similarity 100.0%; Pred. No. 2e-18;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 FAKTCVADESAENCDSLHTLFGDKLC 75
DB 73 FAKTCVADESAENCDSLHTLFGDKLC 99

RESULT 4
Q7YSG3
ID Q7YSG3 PRELIMINARY; PRT; 584 AA.

AC Q7YSG3
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Albumin (Fragment).
GN ALB.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Reininger R., Swoboda I., Bohle B., Hauswirth A.W., Valent P.,
RA Rumpold H., Valenta R., Spitzauer S.;
RT "Escherichia coli expression and purification of recombinant cat
albumin.IgE recognition, induction of basophil activation and
RT lymphoproliferative responses in atopic patients.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ487677; CAD32275.1; -;
FT NON TER 1
SQ SEQUENCE 584 AA; 65908 MW; B51002F12902C9CE CRC64;

Query Match 4.4%; Score 26; DB 6; Length 584;
Best Local Similarity 100.0%; Pred. No. 2.1e-17;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 403 FQNALVRYTKVPQVSTPTLVEVSR 428
DB 403 FQNALVRYTKVPQVSTPTLVEVSR 428

RESULT 5
Q9QVAl
ID Q9QVAl PRELIMINARY; PRT; 21 AA.

AC Q9QVAl
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE 49 kDa protein (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE.
RX MEDLINE=93105636; PubMed=1468220;
RA Zhu L., Crouch R.K.;
RT "Albumin in the cornea is oxidized by hydrogen peroxide."
RA Cornea 11:567-572(1992).

DR HSP; P02768; IE7H.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 1.
SQ SEQUENCE 21 AA; 2429 MW; 26134A3D7CE29FAC CRC64;

Query Match 3.6%; Score 21; DB 11; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.5e-13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGLENFKA 21
DB 1 DAHSEVAHRFKDLGLENFKA 21

RESULT 6
Q8CG74
ID Q8CG74 PRELIMINARY; PRT; 205 AA.

AC Q8CG74
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 23, Last annotation update)
DE Albumin (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvevTACFBr;
RA Van Reeth T., Dreze P.L., Gabant P., Szpirer C., Szpirer J.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ277794; CAC81903.1; -;
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 1.
DR PRINTS; PRO0802; SERUMALBUMIN.
DR SMART; SM00103; ALBUMIN; 1.
FT NON TER 205
SQ SEQUENCE 205 AA; 23624 MW; 0A982E85C7474BF9 CRC64;

Query Match 3.1%; Score 18; DB 11; Length 205;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 KOEPERNECFLOHKDDNP 110
DB 117 KOEPERNECFLOHKDDNP 134

RESULT 7
Q8C7C7
ID Q8C7C7 PRELIMINARY; PRT; 576 AA.

AC Q8C7C7
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Albumin 1 (Fragment).
GN ALB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;

Q7wv14 uncultured
Q7wv13 uncultured
Q7wv12 uncultured
Q7wv11 uncultured
Q7wv10 uncultured
Q7wv9 uncultured
Q7wv8 uncultured
Q7wv7 uncultured
Q7wv6 uncultured
Q7wv5 uncultured
Q7wv4 uncultured

ALIGNMENTS

RESULT 1
Q86YGO PRELIMINARY; PRT; 417 AA.
AC Q86YGO;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to alpha-fetoprotein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041789; AAH41789.1;
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPRO00264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 417 AA; 47360 MW; 16E764033BEF4E8D CRC64;

Query Match 64.1%; Score 375; DB 4; Length 417;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
211 FKAWARLSQRPKAEFEVSKLVTDLTQVTECHGDLLECDRADLAKYICENQDS 270
43 FKAWARLSQRPKAEFEVSKLVTDLTQVTECHGDLLECDRADLAKYICENQDS 102

271 ISSKLECECKPLEKSHCIAEVENDEMPADLPSLAADFVSKDCKYNAKQVFIOMF 330
103 ISSKLECECKPLEKSHCIAEVENDEMPADLPSLAADFVSKDCKYNAKQVFIOMF 162
331 LYVAREHPDYSVLLRLAKTETTTLEKCCAAADPHECYAKVDFEPLVEEPONLIKQ 390
163 LYVAREHPDYSVLLRLAKTETTTLEKCCAAADPHECYAKVDFEPLVEEPONLIKQ 222
391 NCELFEQGYKFNQALLVRYTKVPQVSTPTLVEVSRNLGKVGKCKHPKAMPKCAE 450
223 NCELFEQGYKFNQALLVRYTKVPQVSTPTLVEVSRNLGKVGKCKHPKAMPKCAE 282
451 DYLSVNLQCVLHEKTPVSDRVTKCTESLVNRRPCFSALEVDYVVPKFNATFTFFH 510
283 DYLSVNLQCVLHEKTPVSDRVTKCTESLVNRRPCFSALEVDYVVPKFNATFTFFH 342
511 ADICTLSEKERQIKKQATLVELVKHPKATKEQLKAVMDFFAAVFEKCKKADKCTCPAE 570
343 ADICTLSEKERQIKKQATLVELVKHPKATKEQLKAVMDFFAAVFEKCKKADKCTCPAE 402

QY 571 EGKLVAAQAALGL 585
DB 403 EGKLVAAQAALGL 417

RESULT 2
Q81UK7 PRELIMINARY; PRT; 396 AA.
ID Q81UK7;
AC Q81UK7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to serum albumin precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC035969; AAH35969.1;
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPRO00264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;

Query Match 39.8%; Score 233; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 1.7e-230;
Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 353 YETILEKCCAAADPHECVAKVDFEFKPLVEEPQNLKQNCLEFQELGEYKFNALLVRYT 412
DB 164 YETILEKCCAAADPHECVAKVDFEFKPLVEEPQNLKQNCLEFQELGEYKFNALLVRYT 223
QY 413 KKVPQVSTPTLVEVSRNLGKVGKCKHPKAMPKCAEDYLSVNLQCVLHEKTPVSDR 472
DB 224 KKVPQVSTPTLVEVSRNLGKVGKCKHPKAMPKCAEDYLSVNLQCVLHEKTPVSDR 283
QY 473 VTKCTESLVNRRPCFSALEVDYVVPKFNATFTFHADICTLSEKERQIKKQATLVEL 532
DB 284 VTKCTESLVNRRPCFSALEVDYVVPKFNATFTFHADICTLSEKERQIKKQATLVEL 343
QY 533 VKHKPKATKEQLKAVMDFFAAVFEKCKKADKCTCPAEKGLVAAQAALGL 585
DB 344 VKHKPKATKEQLKAVMDFFAAVFEKCKKADKCTCPAEKGLVAAQAALGL 396

RESULT 3
Q95VB7 PRELIMINARY; PRT; 608 AA.
ID Q95VB7;
AC Q95VB7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Albumin.
OS Schistosoma mansoni (Blood fluke)
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RA Oman A.; Asahi H.; Stadelker M.J.; Loverde P.T.;
RT "Albumin precursor homolog is a novel T helper cell immunogenic egg component in murine infection with Schistosoma mansoni."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
EMBL; AF418550; AAL08579.1; --

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 16:03:04 / Search time 45 Seconds

(without alignment)
4101.737 Million cell updates/sec

Title: US-09-832-929-18

Perfect score: 585

Sequence: 1 DAHKSEVAHFRKDLGEEFNK.....TCFAEGKKLVAASQAALGL 585

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : SPTREMBL 25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	375	64.1	417	4	Q86YG0
2	233	39.8	396	4	Q81UK7
3	27	4.6	608	5	Q95VB7
4	26	4.4	584	6	Q7YSG3
5	21	3.6	21	11	Q9QVAL
6	18	3.1	205	11	Q8CG74
7	18	3.1	576	11	Q8C7C7
8	18	3.1	608	11	Q8C7H3
9	10	1.7	61	6	Q9TS75
10	9	1.5	20	2	Q9RAX7
11	9	1.5	20	11	Q9QUX8
12	8	1.4	20	6	Q9TCZ6
13	8	1.4	44	6	Q95MC2
14	8	1.4	115	12	Q8QML2
15	8	1.4	128	4	Q9HB64
16	8	1.4	273	16	Q9PHJ3

17	8	1.4	289	16	P74538	synchocyst
18	8	1.4	295	16	Q89086	clostridium
19	8	1.4	398	11	Q8ROV5	mus musculus
20	8	1.4	422	16	Q86640	streptomyces
21	8	1.4	468	5	Q8MMD7	drosophila
22	8	1.4	483	17	Q971B5	sulfolobus
23	8	1.4	490	16	Q8XL35	clostridium
24	8	1.4	498	16	Q8FFB6	escherichia
25	8	1.4	509	16	Q92JY9	rhizobium m
26	8	1.4	538	3	Q86729	saccharomyc
27	8	1.4	575	16	Q8XBL3	escherichia
28	8	1.4	575	16	Q83K79	shigella fl
29	8	1.4	595	10	Q85759	oryza sativ
30	8	1.4	602	10	Q857F6	oryza sativ
31	8	1.4	609	6	Q8MJU5	canis famil
32	8	1.4	610	6	Q8MJU6	sus scrofa
33	8	1.4	663	5	Q9VQS3	drosophila
34	8	1.4	663	5	Q8SZ97	drosophila
35	8	1.4	752	11	Q9JL63	mus musculus
36	8	1.4	760	16	Q816K5	pseudomonas
37	8	1.4	948	16	Q7V975	prochloroco
38	8	1.4	1044	10	Q8VY00	arabidopsis
39	8	1.4	1090	10	Q9LQK8	arabidopsis
40	7	1.2	36	16	Q7UL51	rhodospirill
41	7	1.2	63	12	Q9JLK9	hepatitis c
42	7	1.2	65	16	Q81EG4	bacillus ce
43	7	1.2	65	12	Q8530	hepatitis c
44	7	1.2	65	12	Q8525	hepatitis c
45	7	1.2	65	12	Q8524	hepatitis c
46	7	1.2	65	12	Q8521	hepatitis c
47	7	1.2	65	12	Q8522	hepatitis c
48	7	1.2	65	12	Q8523	hepatitis c
49	7	1.2	65	12	Q8531	hepatitis c
50	7	1.2	65	12	Q8532	hepatitis c
51	7	1.2	65	12	Q8527	hepatitis c
52	7	1.2	65	12	Q8528	hepatitis c
53	7	1.2	65	12	Q8526	hepatitis c
54	7	1.2	67	12	Q40646	hepatitis c
55	7	1.2	71	16	Q98A75	rhizobium l
56	7	1.2	73	12	Q9E8X2	hepatitis c
57	7	1.2	73	12	Q9E8X3	hepatitis c
58	7	1.2	73	12	Q9E8X5	hepatitis c
59	7	1.2	73	12	Q9E8W9	hepatitis c
60	7	1.2	73	12	Q9E8X4	hepatitis c
61	7	1.2	73	12	Q9E8X0	hepatitis c
62	7	1.2	73	12	Q9E8W7	hepatitis c
63	7	1.2	73	12	Q9E8X6	hepatitis c
64	7	1.2	73	12	Q9E8W6	hepatitis c
65	7	1.2	73	12	Q9E8W8	hepatitis c
66	7	1.2	73	12	Q9E8X1	hepatitis c
67	7	1.2	74	2	Q7WVK9	uncultured
68	7	1.2	74	2	Q7WVK7	uncultured
69	7	1.2	74	2	Q7WVK6	uncultured
70	7	1.2	74	2	Q7WVK5	uncultured
71	7	1.2	74	2	Q7WVK4	uncultured
72	7	1.2	74	2	Q7WVK3	uncultured
73	7	1.2	74	2	Q7WVK2	uncultured
74	7	1.2	74	2	Q7WVK1	uncultured
75	7	1.2	74	2	Q7WVK0	uncultured
76	7	1.2	74	2	Q7WVJ9	uncultured
77	7	1.2	74	2	Q7WVJ8	uncultured
78	7	1.2	74	2	Q7WVJ7	uncultured
79	7	1.2	74	2	Q7WVJ6	uncultured
80	7	1.2	74	2	Q7WVJ5	uncultured
81	7	1.2	74	2	Q7WVJ4	uncultured
82	7	1.2	74	2	Q7WVJ3	uncultured
83	7	1.2	74	2	Q7WVJ1	uncultured
84	7	1.2	74	2	Q7WVJ0	uncultured
85	7	1.2	74	2	Q7WVI9	uncultured
86	7	1.2	74	2	Q7WVI8	uncultured
87	7	1.2	74	2	Q7WVI7	uncultured
88	7	1.2	74	2	Q7WVI6	uncultured
89	7	1.2	74	2	Q7WVI5	uncultured

QY 541 KEQLKAVMDPFAAFVEKCCADKDKTCFAEGKKLVAAASQAALGL 585
DB 743 KEQLKAVMDPFAAFVEKCCADKDKTCFAEGKKLVAAASQAALGL 787

RESULT 75

US-10-237-667-16
; Sequence 16, Application US/10237667
; Publication No. US20030022308A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Gerard
; Yeh, Parice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESS: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3043
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,667
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993

ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-36,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 787 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 16:

US-10-237-667-16
Query Match 100.0%; Score 585; DB 14; Length 787;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHAFKDLGKLVIAFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
DB 203 DAHKEVAHAFKDLGKLVIAFAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 262
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPFLRLVRPEV 120
DB 263 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPFLRLVRPEV 322
QY 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180

DB 323 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 382
QY 181 KLDELDEGKASSAKORLKCASIQFGERAFKAWAVARLSORFPKRAEFAEYSKLVTDLT 240
DB 383 KLDELDEGKASSAKORLKCASIQFGERAFKAWAVARLSORFPKRAEFAEYSKLVTDLT 442
QY 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 300
DB 443 VHTECCHGDLLECADRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 502
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLVEYARRHPDYSVVLLRLRAKTYETTLEK 360
DB 503 DLPSLAADFVESKDVCKNYAEAKDVLGMFLVEYARRHPDYSVVLLRLRAKTYETTLEK 562
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKVPQVST 420
DB 563 CAAADPHECYAKVDFEFKPLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKVPQVST 622
QY 421 PTLVEVSRLNGYKSKCKHPEAKMPCAEEDYLSVVLNOLCVLHEKTPVSDRVTKCCTES 480
DB 623 PTLVEVSRLNGYKSKCKHPEAKMPCAEEDYLSVVLNOLCVLHEKTPVSDRVTKCCTES 682
QY 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 540
DB 683 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKHKPKAT 742
QY 541 KEQLKAVMDPFAAFVEKCCADKDKTCFAEGKKLVAAASQAALGL 585
DB 743 KEQLKAVMDPFAAFVEKCCADKDKTCFAEGKKLVAAASQAALGL 787

Search completed: April 19, 2004, 16:13:18
Job time : 54 secs

/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: HSA-HIL-11
US-10-609-346-2

Query Match 100.0%; Score 585; DB 12; Length 763;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLEENFKALVLIAPAOYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHRFKDLEENFKALVLIAPAOYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGENADCCAKQBERNECFLOHKDDNPNLRLVRREV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGENADCCAKQBERNECFLOHKDDNPNLRLVRREV 120

QY 121 DVMTAFPHDNEETFLKYLVEIARRHPFYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 121 DVMTAFPHDNEETFLKYLVEIARRHPFYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 180

QY 181 KLDELURDEGKASSAKQRLKCSLOKGERAFKAWAVARLSQRPKAEFAEVSCLVTDLTK 240
Db 181 KLDELURDEGKASSAKQRLKCSLOKGERAFKAWAVARLSQRPKAEFAEVSCLVTDLTK 240

QY 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKCECCEKPLEKSHGICIAEVENDEMPA 300
Db 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKCECCEKPLEKSHGICIAEVENDEMPA 300

QY 301 DLPSLAADPVESKOVCKNYAEAKDVFLGMFLYIARRHPDYSVLLLRLLAKTYETTLK 360
Db 301 DLPSLAADPVESKOVCKNYAEAKDVFLGMFLYIARRHPDYSVLLLRLLAKTYETTLK 360

QY 361 CAAADPHCYAKVDFEKPFLVEEPQNLKQNCLEFQGLGEYKFQNALVRYTKVPQVST 420
Db 361 CAAADPHCYAKVDFEKPFLVEEPQNLKQNCLEFQGLGEYKFQNALVRYTKVPQVST 420

QY 421 PTLVEVSNLKGVSKCKCKPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSNLKGVSKCKCKPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRVTKCCTES 480

QY 481 LVNRRPCFSALEVDVETVYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVGHKPKAT 540
Db 481 LVNRRPCFSALEVDVETVYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVGHKPKAT 540

QY 541 KEQLKAVMDDFAAAFVEKCKADDKCTCFABEGKKLVAAASQAALGL 585
Db 541 KEQLKAVMDDFAAAFVEKCKADDKCTCFABEGKKLVAAASQAALGL 585

RESULT 74
US-09-984-186-16
Sequence 16, Application US/09984186
Patent No. US20020151011A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3043
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/984,186
FILING DATE: 29-Oct-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 787 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-984-186-16

Query Match 100.0%; Score 585; DB 9; Length 787;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLEENFKALVLIAPAOYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60
Db 203 DAHSEVAHRFKDLEENFKALVLIAPAOYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 262

QY 61 NCDKSLHTLFGDKLCTVATLRETYGENADCCAKQBERNECFLOHKDDNPNLRLVRREV 120
Db 263 NCDKSLHTLFGDKLCTVATLRETYGENADCCAKQBERNECFLOHKDDNPNLRLVRREV 322

QY 121 DVMTAFPHDNEETFLKYLVEIARRHPFYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 323 DVMTAFPHDNEETFLKYLVEIARRHPFYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 382

QY 181 KLDELURDEGKASSAKQRLKCSLOKGERAFKAWAVARLSQRPKAEFAEVSCLVTDLTK 240
Db 383 KLDELURDEGKASSAKQRLKCSLOKGERAFKAWAVARLSQRPKAEFAEVSCLVTDLTK 442

QY 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKCECCEKPLEKSHGICIAEVENDEMPA 300
Db 443 VHTECHGDLLECADRADLAKYICENQDSISSKLKCECCEKPLEKSHGICIAEVENDEMPA 502

QY 301 DLPSLAADPVESKOVCKNYAEAKDVFLGMFLYIARRHPDYSVLLLRLLAKTYETTLK 360
Db 503 DLPSLAADPVESKOVCKNYAEAKDVFLGMFLYIARRHPDYSVLLLRLLAKTYETTLK 562

QY 361 CAAADPHCYAKVDFEKPFLVEEPQNLKQNCLEFQGLGEYKFQNALVRYTKVPQVST 420
Db 563 CAAADPHCYAKVDFEKPFLVEEPQNLKQNCLEFQGLGEYKFQNALVRYTKVPQVST 622

QY 421 PTLVEVSNLKGVSKCKCKPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRVTKCCTES 480
Db 623 PTLVEVSNLKGVSKCKCKPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRVTKCCTES 682

QY 481 LVNRRPCFSALEVDVETVYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVGHKPKAT 540
Db 683 LVNRRPCFSALEVDVETVYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVGHKPKAT 742

; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 689
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-33

Query Match 100.0%; Score 585; DB 12; Length 689;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
DB 2 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 61
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 62 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 121
QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
DB 122 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 181
QY 181 KLDELDEGKASSAKQRLKASLOKGERAFKAWAVARLSORPPKAEFAVSKLVTDLTK 240
DB 182 KLDELDEGKASSAKQRLKASLOKGERAFKAWAVARLSORPPKAEFAVSKLVTDLTK 241
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILLEKSHCIAEVENDEMPA 300
DB 242 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILLEKSHCIAEVENDEMPA 301
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYIYARRHPDYSVLLRLAKTYETTTLEKC 360
DB 302 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYIYARRHPDYSVLLRLAKTYETTTLEKC 361
QY 361 CAADDPHECVAKVDFEFLVEEPONLIKONCELFEOLGEYKFNALLVRYTKVPQVST 420
DB 362 CAADDPHECVAKVDFEFLVEEPONLIKONCELFEOLGEYKFNALLVRYTKVPQVST 421
QY 421 PTLVEVSRNLGVSKCKGHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 422 PTLVEVSRNLGVSKCKGHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 481
QY 481 LVNRRPCFSALEVDYVYKPEFNAETTFHADICTLSEKERQIKQTALVELVKHKPKAT 540
DB 482 LVNRRPCFSALEVDYVYKPEFNAETTFHADICTLSEKERQIKQTALVELVKHKPKAT 541
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDDKCTCFABEGKKLVAASQAALGL 585
DB 542 KEQLKAVMDDDFAAFVEKCKCKADDDKCTCFABEGKKLVAASQAALGL 586

RESULT 72
US-10-233-675A-13
; Sequence 13, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04

; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 689
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion protein human abrogen
US-10-233-675A-13

Query Match 100.0%; Score 585; DB 15; Length 689;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
DB 2 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 61
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 62 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 121
QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
DB 122 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 181
QY 181 KLDELDEGKASSAKQRLKASLOKGERAFKAWAVARLSORPPKAEFAVSKLVTDLTK 240
DB 182 KLDELDEGKASSAKQRLKASLOKGERAFKAWAVARLSORPPKAEFAVSKLVTDLTK 241
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILLEKSHCIAEVENDEMPA 300
DB 242 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILLEKSHCIAEVENDEMPA 301
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYIYARRHPDYSVLLRLAKTYETTTLEKC 360
DB 302 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYIYARRHPDYSVLLRLAKTYETTTLEKC 361
QY 361 CAADDPHECVAKVDFEFLVEEPONLIKONCELFEOLGEYKFNALLVRYTKVPQVST 420
DB 362 CAADDPHECVAKVDFEFLVEEPONLIKONCELFEOLGEYKFNALLVRYTKVPQVST 421
QY 421 PTLVEVSRNLGVSKCKGHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 422 PTLVEVSRNLGVSKCKGHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 481
QY 481 LVNRRPCFSALEVDYVYKPEFNAETTFHADICTLSEKERQIKQTALVELVKHKPKAT 540
DB 482 LVNRRPCFSALEVDYVYKPEFNAETTFHADICTLSEKERQIKQTALVELVKHKPKAT 541
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDDKCTCFABEGKKLVAASQAALGL 585
DB 542 KEQLKAVMDDDFAAFVEKCKCKADDDKCTCFABEGKKLVAASQAALGL 586

RESULT 73
US-10-609-346-2
; Sequence 2, Application US/10609346
; Publication No. US20040063635A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Zailin
; APPLICANT: Fu, Yan
; TITLE OF INVENTION: RECOMBINANT HUMAN ALBUMIN FUSION PROTEINS WITH LONG-LASTING BIO
; TITLE OF INVENTION: EFFECTS
; FILE REFERENCE: ZYU-0603
; CURRENT APPLICATION NUMBER: US/10/609,346
; CURRENT FILING DATE: 2003-06-26
; PRIOR APPLICATION NUMBER: US 60/392,948
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 763

APPLICANT: Fong, Timothy
APPLICANT: Brockstedt, Dirk
TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
TITLE OF INVENTION: Them To Inhibit Angiogenesis
FILE REFERENCE: ST01027
CURRENT APPLICATION NUMBER: US/10/233,675A
CURRENT FILING DATE: 2002-09-04
PRIOR APPLICATION NUMBER: 60/316,300
PRIOR FILING DATE: 2001-09-04
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn version 3.1
SEQ ID NO 18
LENGTH: 688
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: fusion protein human abrogen
US-10-233-675A-18

Query Match 100.0%; Score 585; DB 15; Length 688;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DAHKEVAHFPKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHKEVAHFPKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEWADCCAKQEPERNECFLOHKDNDPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEWADCCAKQEPERNECFLOHKDNDPNLRLVRPEV 120
Qy 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Db 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Qy 181 KLDLRDEGKASSAKQRLKASLOKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240
Db 181 KLDLRDEGKASSAKQRLKASLOKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240
Qy 241 VHTCCCHGDLLECCADRDADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 300
Db 241 VHTCCCHGDLLECCADRDADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 300
Qy 301 DLPSLAADFVSKDVCNKYAEAKDVLGMFLYAEARRHPDYSVLLRLAKTYETTLK 360
Db 301 DLPSLAADFVSKDVCNKYAEAKDVLGMFLYAEARRHPDYSVLLRLAKTYETTLK 360
Qy 361 CAADDPHECYAKVDFEFPKLVPEEPQNLKONCELFQELGEYKFNALLVRYTKVPQVST 420
Db 361 CAADDPHECYAKVDFEFPKLVPEEPQNLKONCELFQELGEYKFNALLVRYTKVPQVST 420
Qy 421 PTLVEVSRLKGVSKCKKHPEAKMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRLKGVSKCKKHPEAKMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Qy 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKERQIKKOTALVELVGHKPKAT 540
Db 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKERQIKKOTALVELVGHKPKAT 540
Qy 541 KEQLKAVMDPFAAFVEKCKCKADKCTCFABEGKKLVAASQAALGL 585
Db 541 KEQLKAVMDPFAAFVEKCKCKADKCTCFABEGKKLVAASQAALGL 585

RESULT 70
US-10-424-999-13
Sequence 13, Application US/10424999
Publication No. US20040052810A1
GENERAL INFORMATION:
APPLICANT: Nesbit, Mark
APPLICANT: Cameron, Beatrice
APPLICANT: Blanche, Francis
TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for

TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
FILE REFERENCE: ST01027-A
CURRENT APPLICATION NUMBER: US/10/424,999
CURRENT FILING DATE: 2003-04-29
PRIOR APPLICATION NUMBER: 10/233,675
PRIOR FILING DATE: 2002-09-04
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn version 3.2
SEQ ID NO 13
LENGTH: 689
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-13

Query Match 100.0%; Score 585; DB 12; Length 689;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DAHKEVAHFPKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 2 DAHKEVAHFPKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 61
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEWADCCAKQEPERNECFLOHKDNDPNLRLVRPEV 120
Db 62 NCDKSLHTLFGDKLCTVATLRETYGEWADCCAKQEPERNECFLOHKDNDPNLRLVRPEV 121
Qy 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Db 122 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 181
Qy 181 KLDLRDEGKASSAKQRLKASLOKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240
Db 182 KLDLRDEGKASSAKQRLKASLOKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 241
Qy 241 VHTCCCHGDLLECCADRDADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 300
Db 242 VHTCCCHGDLLECCADRDADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 301
Qy 301 DLPSLAADFVSKDVCNKYAEAKDVLGMFLYAEARRHPDYSVLLRLAKTYETTLK 360
Db 302 DLPSLAADFVSKDVCNKYAEAKDVLGMFLYAEARRHPDYSVLLRLAKTYETTLK 361
Qy 361 CAADDPHECYAKVDFEFPKLVPEEPQNLKONCELFQELGEYKFNALLVRYTKVPQVST 420
Db 362 CAADDPHECYAKVDFEFPKLVPEEPQNLKONCELFQELGEYKFNALLVRYTKVPQVST 421
Qy 421 PTLVEVSRLKGVSKCKKHPEAKMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 422 PTLVEVSRLKGVSKCKKHPEAKMPCAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES 481
Qy 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKERQIKKOTALVELVGHKPKAT 540
Db 482 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKERQIKKOTALVELVGHKPKAT 541
Qy 541 KEQLKAVMDPFAAFVEKCKCKADKCTCFABEGKKLVAASQAALGL 585
Db 542 KEQLKAVMDPFAAFVEKCKCKADKCTCFABEGKKLVAASQAALGL 586

RESULT 71
US-10-425-000-33
Sequence 33, Application US/10425000
Publication No. US20040052777A1
GENERAL INFORMATION:
APPLICANT: Nesbit, Mark
APPLICANT: Cameron, Beatrice
APPLICANT: Blanche, Francis
TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
TITLE OF INVENTION: Angiogenesis
FILE REFERENCE: ST01027-B
CURRENT APPLICATION NUMBER: US/10/425,000

RESULT 67
 US-10-424-999-18
 ; Sequence 18, Application US/10424999
 ; Publication No. US20040052810A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Nesbit, Mark
 ; APPLICANT: Cameron, Beatrice
 ; APPLICANT: Blanche, Francis
 ; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
 ; FILE REFERENCE: ST01027-A
 ; CURRENT APPLICATION NUMBER: US/10/424,999
 ; PRIOR FILING DATE: 2003-04-29
 ; PRIOR FILING DATE: 2002-09-04
 ; NUMBER OF SEQ ID NOS: 70
 ; SOFTWARE: Patent in version 3.2
 ; SEQ ID NO 18
 ; TYPE: PRT
 ; LENGTH: 688
 ; ORGANISM: Artificial Sequence
 ; OTHER INFORMATION: Fusion protein human abrogen
 US-10-424-999-18

 Query Match 100.0%; Score 585; DB 12; Length 688;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 DAHSEVAHFRFDLGEENFKALVLIATAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
 Db 1 DAHSEVAHFRFDLGEENFKALVLIATAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60

 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

 QY 121 DVMTAFHNEETFLKYLVEIARRHPFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
 Db 121 DVMTAFHNEETFLKYLVEIARRHPFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180

 QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
 Db 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240

 QY 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECCXPLEKSHCIAEVENDEMPA 300
 Db 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECCXPLEKSHCIAEVENDEMPA 300

 QY 301 DLPSLAADFVSKDVCNKYAEAKDVLGMFLYEYARRHPDYSVLLLRLLAKTYETTLK 360
 Db 301 DLPSLAADFVSKDVCNKYAEAKDVLGMFLYEYARRHPDYSVLLLRLLAKTYETTLK 360

 QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLKQNCLEFQOLGEYKFNQALLVRYTKKVPQVST 420
 Db 361 CAAADPHECYAKVDFEFKPLVEEPQNLKQNCLEFQOLGEYKFNQALLVRYTKKVPQVST 420

 QY 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVNLQCLVHEKTPVSDRVTKCCTES 480
 Db 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVNLQCLVHEKTPVSDRVTKCCTES 480

 QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540
 Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540

 QY 541 KEQLKAVMDDFAAAFVEKCKKADDKETCFABEGKCLVAASQAALGL 585
 Db 541 KEQLKAVMDDFAAAFVEKCKKADDKETCFABEGKCLVAASQAALGL 585

RESULT 68
 US-10-425-000-38
 ; Sequence 38, Application US/10425000

; Publication No. US20040052777A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Nesbit, Mark
 ; APPLICANT: Cameron, Beatrice
 ; APPLICANT: Blanche, Francis
 ; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit Angiogenesis
 ; FILE REFERENCE: ST01027-B
 ; CURRENT APPLICATION NUMBER: US/10/425,000
 ; PRIOR FILING DATE: 2003-04-29
 ; PRIOR FILING DATE: 2002-09-04
 ; NUMBER OF SEQ ID NOS: 105
 ; SOFTWARE: Patent in version 3.2
 ; SEQ ID NO 38
 ; LENGTH: 688
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; OTHER INFORMATION: Human derived fusion protein
 US-10-425-000-38

 Query Match 100.0%; Score 585; DB 12; Length 688;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 DAHSEVAHFRFDLGEENFKALVLIATAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
 Db 1 DAHSEVAHFRFDLGEENFKALVLIATAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60

 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

 QY 121 DVMTAFHNEETFLKYLVEIARRHPFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
 Db 121 DVMTAFHNEETFLKYLVEIARRHPFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180

 QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
 Db 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240

 QY 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECCXPLEKSHCIAEVENDEMPA 300
 Db 241 VHTECHGDLLECADRADLAKYICENQDSISSKLKECCXPLEKSHCIAEVENDEMPA 300

 QY 301 DLPSLAADFVSKDVCNKYAEAKDVLGMFLYEYARRHPDYSVLLLRLLAKTYETTLK 360
 Db 301 DLPSLAADFVSKDVCNKYAEAKDVLGMFLYEYARRHPDYSVLLLRLLAKTYETTLK 360

 QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLKQNCLEFQOLGEYKFNQALLVRYTKKVPQVST 420
 Db 361 CAAADPHECYAKVDFEFKPLVEEPQNLKQNCLEFQOLGEYKFNQALLVRYTKKVPQVST 420

 QY 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVNLQCLVHEKTPVSDRVTKCCTES 480
 Db 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVNLQCLVHEKTPVSDRVTKCCTES 480

 QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540
 Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540

 QY 541 KEQLKAVMDDFAAAFVEKCKKADDKETCFABEGKCLVAASQAALGL 585
 Db 541 KEQLKAVMDDFAAAFVEKCKKADDKETCFABEGKCLVAASQAALGL 585

RESULT 69
 US-10-233-675A-18
 ; Sequence 18, Application US/10233675A
 ; Publication No. US2003022898A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Nesbit, Mark

Db 643 KQQLKAVNMDDFAAAFVEKCCCKADDDKTCFABEGKKLVAASQAALGL 687

RESULT 66
US-10-233-675A-17
; Sequence 17, Application US/10233675A
; Publication No. US2003022829A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Arogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion protein human abrogen
US-10-233-675A-17

Query Match 100.0%; Score 585; DB 15; Length 687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	DAKSEVAHRPKDILGEENFKALVLIARAOYLQCCPDHVKLVNVEYFAKTCVADESSE	60
Db	103	DAKSEVAHRPKDILGEENFKALVLIARAOYLQCCPDHVKLVNVEYFAKTCVADESSE	162
QY	61	NCCKSLHTLFGDKLCTVATLRETTGEMADCAKQEPNERNECFLOHKDDNPMLRLVRPEV	120
Db	163	NCCKSLHTLFGDKLCTVATLRETTGEMADCAKQEPNERNECFLOHKDDNPMLRLVRPEV	222
QY	121	DVWCATFDHNEETFLKKLYLIARHRYFYFAPELLFFAKYKAAFTTECCQAADKKAACLLP	180
Db	223	DVWCATFDHNEETFLKKLYLIARHRYFYFAPELLFFAKYKAAFTTECCQAADKKAACLLP	282
QY	181	KLDELURDEGKASSAKORUKCASLOKFGERAFKAWAVARLSORPKAFRAVSKLVTDLTK	240
Db	283	KLDELURDEGKASSAKORUKCASLOKFGERAFKAWAVARLSORPKAFRAVSKLVTDLTK	342
QY	241	VHTECHGDLLECADRRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVNDEMPA	300
Db	343	VHTECHGDLLECADRRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVNDEMPA	402
QY	301	DLPSLAADPVESKDVCKNYAKADKDFLGMFLYEYARRHPDYSVVLLRLAKTVEITLKC	360
Db	403	DLPSLAADPVESKDVCKNYAKADKDFLGMFLYEYARRHPDYSVVLLRLAKTVEITLKC	462
QY	361	CAAADPHECYAKVDFEPPKPLVEEPPNLIKQNCLELFEQIGYKQFNALLVRYTKKVPQVST	420
Db	463	CAAADPHECYAKVDFEPPKPLVEEPPNLIKQNCLELFEQIGYKQFNALLVRYTKKVPQVST	522
QY	421	PTLVEVSRNLGVGSKCKCKPEAKMPCAEDYLSVNLQNLVHLHEKTPVSDRVTKCCTES	480
Db	523	PTLVEVSRNLGVGSKCKCKPEAKMPCAEDYLSVNLQNLVHLHEKTPVSDRVTKCCTES	582
QY	481	LNVNRPCPSALVEDVTYYPKEFNATETFEHADICTLSEKEROIKKOTALVELVKKHPKAT	540
Db	583	LNVNRPCPSALVEDVTYYPKEFNATETFEHADICTLSEKEROIKKOTALVELVKKHPKAT	642
QY	541	KQQLKAVNMDDFAAAFVEKCCCKADDDKTCFABEGKKLVAASQAALGL	585
Db	643	KQQLKAVNMDDFAAAFVEKCCCKADDDKTCFABEGKKLVAASQAALGL	687

Qy	541	KEQLKAVMDDFAAAFVEKCKKADDDKETCFABEGKGLVAASQAALGL	585
Db	643	KEQLKAVMDDFAAAFVEKCKKADDDKETCFABEGKGLVAASQAALGL	687
RESULT 65			
US-10-425-000-37			
; Sequence 37, Application US/10425000			
; Publication No. US20040052777A1			
; GENERAL INFORMATION:			
; APPLICANT: Nesbitt, Mark			
; APPLICANT: Cameron, Beatrice			
; APPLICANT: Blanche, Francis			
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit			
; TITLE OF INVENTION: Angiogenesis			
; FILE REFERENCE: ST01027-B			
; CURRENT APPLICATION NUMBER: US/10/425,000			
; CURRENT FILING DATE: 2003-04-29			
; PRIOR APPLICATION NUMBER: 10/233,675			
; PRIOR FILING DATE: 2002-09-04			
; NUMBER OF SEQ ID NOS: 105			
; SOFTWARE: Patent version 3.2			
; SEQ ID NO 37			
; LENGTH: 687			
; TYPE: PRT			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: Human derived fusion protein			
US-10-425-000-37			
Query Match 100.0%; Score 585; DB 12; Length 687;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	DAHKEVAHRPKDGLGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEPAKTCVADESAB	60
Db	103	DAHKEVAHRPKDGLGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEPAKTCVADESAB	162
Qy	61	NDCKSLHTLFGDKLTCTVATLTREYGEADCAKQEPERNECFLOHKDDNPRLVRPEV	120
Db	163	NDCKSLHTLFGDKLTCTVATLTREYGEADCAKQEPERNECFLOHKDDNPRLVRPEV	222
Qy	121	DVMCTAFHDNEETFLKCYLIARHPFYFAPELFFAKYKAAFTCCQAADKAAACLLP	180
Db	223	DVMCTAFHDNEETFLKCYLIARHPFYFAPELFFAKYKAAFTCCQAADKAAACLLP	282
Qy	181	KLDELRDEGKASSAKORLKCASLOKQGERAFKANAVARLSORPKPAFPAVSKLVTDLTN	240
Db	283	KLDELRDEGKASSAKORLKCASLOKQGERAFKANAVARLSORPKPAFPAVSKLVTDLTN	342
Qy	241	VHTECGHDLLECADDRLADLAKYCENQDISSSKLKCECKPLLEKSHCIAEVENDEMA	300
Db	343	VHTECGHDLLECADDRLADLAKYCENQDISSSKLKCECKPLLEKSHCIAEVENDEMA	402
Qy	301	DLPSLAADPFVESKDVCKYAEAKDVLGMFLVEYARRHPDYSVLLRLAKVTETTLEKC	360
Db	403	DLPSLAADPFVESKDVCKYAEAKDVLGMFLVEYARRHPDYSVLLRLAKVTETTLEKC	462
Qy	361	CAAADPHECYAKVFDEFKPLVEEPONLTKQNCLEFEOLGKYKFNALLVRYTKKVPQVST	420
Db	463	CAAADPHECYAKVFDEFKPLVEEPONLTKQNCLEFEOLGKYKFNALLVRYTKKVPQVST	522
Qy	421	PTLVEVSRNLGKVGSKCKKHPEAKMPCADBYLSVNLNQLCVLHEKTFPSDRVTCKCTES	480
Db	523	PTLVEVSRNLGKVGSKCKKHPEAKMPCADBYLSVNLNQLCVLHEKTFPSDRVTCKCTES	582
Qy	481	LNVNRPCFSALVEDEITYPKPENAETFFHADICTLSEKERQIKQOTALVELVKKHPKAT	540
Db	583	LNVNRPCFSALVEDEITYPKPENAETFFHADICTLSEKERQIKQOTALVELVKKHPKAT	642
Qy	541	KEQLKAVMDDFAAAFVEKCKKADDDKETCFABEGKGLVAASQAALGL	585

QY 361 CAAADPHECYAKVDFEFPKPLVEBPQNLIKQNCBFLQGLGEYFQNALVRYTKKVPQVST 420
Db 456 CAAADPHECYAKVDFEFPKPLVEBPQNLIKQNCBFLQGLGEYFQNALVRYTKKVPQVST 515
QY 421 PTLVEVSRNLGKVGSKCKPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 516 PTLVEVSRNLGKVGSKCKPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 575
QY 481 LVNRRPCFSALEVDETVVPKEFNAETFFHADICTLSEKERQIKKQATLVELVGHKPKAT 540
Db 576 LVNRRPCFSALEVDETVVPKEFNAETFFHADICTLSEKERQIKKQATLVELVGHKPKAT 635
QY 541 KEQLKAVMDPFAAFVEKCKKADDDKTCFAEGKKLVAASQAALGL 585
Db 636 KEQLKAVMDPFAAFVEKCKKADDDKTCFAEGKKLVAASQAALGL 680

RESULT 63

US-10-153-604A-123
; Sequence 123, Application US/10153604A
; Publication No. US20030143191A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,604A
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 123
; LENGTH: 680
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-604A-123

Query Match 100.0%; Score 585; DB 14; Length 680;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHRFKDLGEENFKALVLIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 96 DAHSEVAHRFKDLGEENFKALVLIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 155
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
Db 156 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 215
QY 121 DVNCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
Db 216 DVNCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 275
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
Db 276 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 335
QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILKSHCHIAEVENDEMPA 300
Db 336 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILKSHCHIAEVENDEMPA 395
QY 301 DLPSLAADFVSKDVCNKYAEAKDVFLGMFLYEYARRHPDYSVVLLLRALKTYETTTLEKC 360
Db 396 DLPSLAADFVSKDVCNKYAEAKDVFLGMFLYEYARRHPDYSVVLLLRALKTYETTTLEKC 455
QY 361 CAAADPHECYAKVDFEFPKPLVEBPQNLIKQNCBFLQGLGEYFQNALVRYTKKVPQVST 420
Db 456 CAAADPHECYAKVDFEFPKPLVEBPQNLIKQNCBFLQGLGEYFQNALVRYTKKVPQVST 515
QY 421 PTLVEVSRNLGKVGSKCKPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 516 PTLVEVSRNLGKVGSKCKPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 575

QY 481 LVNRRPCFSALEVDETVVPKEFNAETFFHADICTLSEKERQIKKQATLVELVGHKPKAT 540
Db 576 LVNRRPCFSALEVDETVVPKEFNAETFFHADICTLSEKERQIKKQATLVELVGHKPKAT 635
QY 541 KEQLKAVMDPFAAFVEKCKKADDDKTCFAEGKKLVAASQAALGL 585
Db 636 KEQLKAVMDPFAAFVEKCKKADDDKTCFAEGKKLVAASQAALGL 680
RESULT 64
US-10-424-999-17
; Sequence 17, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 17
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-17

Query Match 100.0%; Score 585; DB 12; Length 687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHRFKDLGEENFKALVLIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 103 DAHSEVAHRFKDLGEENFKALVLIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 162
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
Db 163 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 222
QY 121 DVNCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
Db 223 DVNCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 282
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
Db 283 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 342
QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILKSHCHIAEVENDEMPA 300
Db 343 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILKSHCHIAEVENDEMPA 402
QY 301 DLPSLAADFVSKDVCNKYAEAKDVFLGMFLYEYARRHPDYSVVLLLRALKTYETTTLEKC 360
Db 403 DLPSLAADFVSKDVCNKYAEAKDVFLGMFLYEYARRHPDYSVVLLLRALKTYETTTLEKC 462
QY 361 CAAADPHECYAKVDFEFPKPLVEBPQNLIKQNCBFLQGLGEYFQNALVRYTKKVPQVST 420
Db 463 CAAADPHECYAKVDFEFPKPLVEBPQNLIKQNCBFLQGLGEYFQNALVRYTKKVPQVST 522
QY 421 PTLVEVSRNLGKVGSKCKPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 523 PTLVEVSRNLGKVGSKCKPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 582
QY 481 LVNRRPCFSALEVDETVVPKEFNAETFFHADICTLSEKERQIKKQATLVELVGHKPKAT 540
Db 583 LVNRRPCFSALEVDETVVPKEFNAETFFHADICTLSEKERQIKKQATLVELVGHKPKAT 642

QY 121 DVNCTAFHDNEBETFLKYLVEIARRHPFYFAPPELLFFAKRYKAAETCCQADKAACLLP 180
 DB 213 DVNCTAFHDNEBETFLKYLVEIARRHPFYFAPPELLFFAKRYKAAETCCQADKAACLLP 272
 QY 181 KLDLRDGGKASSAKQRLKCSAQKQGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 240
 DB 273 KLDLRDGGKASSAKQRLKCSAQKQGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 332
 QY 241 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKJKECCCKPILLESKSHCIAEVENDEMPA 300
 DB 333 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKJKECCCKPILLESKSHCIAEVENDEMPA 392
 QY 301 DLPSLAADFVSKDVCKVKNYAEAKDVFLGMFLYIYARRHPDYSVVLRLRLAKTYETTLK 360
 DB 393 DLPSLAADFVSKDVCKVKNYAEAKDVFLGMFLYIYARRHPDYSVVLRLRLAKTYETTLK 452
 QY 361 CAADPHCYAKVDFPKPLVEBPONLIKONCELPOLGEYKFQNALVRYTKVPOYST 420
 DB 453 CAADPHCYAKVDFPKPLVEBPONLIKONCELPOLGEYKFQNALVRYTKVPOYST 512
 QY 421 PTLVEVSRLKGVSKCKCKHPEAKRMPCAEDYLSVVLNOLCVLHKEKTPVSDRVTKCCTES 480
 DB 513 PTLVEVSRLKGVSKCKCKHPEAKRMPCAEDYLSVVLNOLCVLHKEKTPVSDRVTKCCTES 572
 QY 481 LVNRRPCFSALEVDVETVPKFEFNAETFTFHADICTLSEKEROIKKOTALVELVXHKPKAT 540
 DB 573 LVNRRPCFSALEVDVETVPKFEFNAETFTFHADICTLSEKEROIKKOTALVELVXHKPKAT 632
 QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585
 DB 633 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 677
 RESULT 61
 US-10-153-604A-125
 ; Sequence 125, Application US/10153604A
 ; Publication No. US20030143191A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bell et al.
 ; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
 ; FILE REFERENCE: PF556
 ; CURRENT APPLICATION NUMBER: US/10/153,604A
 ; PRIOR FILING DATE: 2002-05-24
 ; PRIOR APPLICATION NUMBER: 60/293,212
 ; NUMBER OF SEQ ID NOS: 137
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 125
 ; LENGTH: 677
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-153-604A-125

Query Match 100.0%; Score 585; DB 14; Length 677;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAQYLQCCPPEDHVKLVNEVTEPAKTCVADESAAE 60
 DB 93 DAHKSEVAHRFKDLGEENFKALVLIAPAQYLQCCPPEDHVKLVNEVTEPAKTCVADESAAE 152
 QY 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQBERNECFLOHKKDNPFLRLVRREV 120
 DB 153 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQBERNECFLOHKKDNPFLRLVRREV 212
 QY 121 DVNCTAFHDNEBETFLKYLVEIARRHPFYFAPPELLFFAKRYKAAETCCQADKAACLLP 180
 DB 213 DVNCTAFHDNEBETFLKYLVEIARRHPFYFAPPELLFFAKRYKAAETCCQADKAACLLP 272
 QY 181 KLDLRDGGKASSAKQRLKCSAQKQGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 240
 DB 273 KLDLRDGGKASSAKQRLKCSAQKQGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 332

QY 241 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKJKECCCKPILLESKSHCIAEVENDEMPA 300
 DB 333 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKJKECCCKPILLESKSHCIAEVENDEMPA 392
 QY 301 DLPSLAADFVSKDVCKVKNYAEAKDVFLGMFLYIYARRHPDYSVVLRLRLAKTYETTLK 360
 DB 393 DLPSLAADFVSKDVCKVKNYAEAKDVFLGMFLYIYARRHPDYSVVLRLRLAKTYETTLK 452
 QY 361 CAADPHCYAKVDFPKPLVEBPONLIKONCELPOLGEYKFQNALVRYTKVPOYST 420
 DB 453 CAADPHCYAKVDFPKPLVEBPONLIKONCELPOLGEYKFQNALVRYTKVPOYST 512
 QY 421 PTLVEVSRLKGVSKCKCKHPEAKRMPCAEDYLSVVLNOLCVLHKEKTPVSDRVTKCCTES 480
 DB 513 PTLVEVSRLKGVSKCKCKHPEAKRMPCAEDYLSVVLNOLCVLHKEKTPVSDRVTKCCTES 572
 QY 481 LVNRRPCFSALEVDVETVPKFEFNAETFTFHADICTLSEKEROIKKOTALVELVXHKPKAT 540
 DB 573 LVNRRPCFSALEVDVETVPKFEFNAETFTFHADICTLSEKEROIKKOTALVELVXHKPKAT 632
 QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585
 DB 633 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 677
 RESULT 62
 US-10-153-064-123
 ; Sequence 123, Application US/10153064
 ; Publication No. US20020142814A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bell et al.
 ; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
 ; FILE REFERENCE: PF556
 ; CURRENT APPLICATION NUMBER: US/10/153,064
 ; PRIOR FILING DATE: 2002-05-24
 ; PRIOR APPLICATION NUMBER: 60/293,212
 ; NUMBER OF SEQ ID NOS: 137
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 123
 ; LENGTH: 680
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-153-064-123
 Query Match 100.0%; Score 585; DB 13; Length 680;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAQYLQCCPPEDHVKLVNEVTEPAKTCVADESAAE 60
 DB 96 DAHKSEVAHRFKDLGEENFKALVLIAPAQYLQCCPPEDHVKLVNEVTEPAKTCVADESAAE 155
 QY 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQBERNECFLOHKKDNPFLRLVRREV 120
 DB 156 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQBERNECFLOHKKDNPFLRLVRREV 215
 QY 121 DVNCTAFHDNEBETFLKYLVEIARRHPFYFAPPELLFFAKRYKAAETCCQADKAACLLP 180
 DB 216 DVNCTAFHDNEBETFLKYLVEIARRHPFYFAPPELLFFAKRYKAAETCCQADKAACLLP 275
 QY 181 KLDLRDGGKASSAKQRLKCSAQKQGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 240
 DB 276 KLDLRDGGKASSAKQRLKCSAQKQGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 335
 QY 241 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKJKECCCKPILLESKSHCIAEVENDEMPA 300
 DB 336 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKJKECCCKPILLESKSHCIAEVENDEMPA 395
 QY 301 DLPSLAADFVSKDVCKVKNYAEAKDVFLGMFLYIYARRHPDYSVVLRLRLAKTYETTLK 360
 DB 396 DLPSLAADFVSKDVCKVKNYAEAKDVFLGMFLYIYARRHPDYSVVLRLRLAKTYETTLK 455

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-604A-127

Query Match      100.0%; Score 585; DB 14; Length 676;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 92 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 151

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 120
DB 152 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 211

QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 212 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 271

QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAEVSKLVDLTG 240
DB 272 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAEVSKLVDLTG 331

QY 241 VHTCCGHDLLCADDADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
DB 332 VHTCCGHDLLCADDADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 391

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVVLLRLAKTYETTLEKC 360
DB 392 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVVLLRLAKTYETTLEKC 451

QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNQALLVRYTKVPQVST 420
DB 452 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNQALLVRYTKVPQVST 511

QY 421 PTLVEVSRNLKGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 512 PTLVEVSRNLKGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 571

QY 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVXHKPKAT 540
DB 572 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVXHKPKAT 631

QY 541 KEQLKAVMDDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
DB 632 KEQLKAVMDDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 676

RESULT 59
US-10-153-604A-129
; Sequence 129, Application US/10153604A
; Publication No. US20030143191A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 129
; LENGTH: 676
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-604A-129

Query Match      100.0%; Score 585; DB 14; Length 676;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 93 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 152

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 120
DB 153 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 212
```

RESULT 56
US-10-153-064-127
; Sequence 127, Application US/10153064
; Publication No. US20020142814A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 127
; LENGTH: 676
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-127

Query Match 100.0%; Score 585; DB 13; Length 676;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRPKDGLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 92 DAHKEVAHRPKDGLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 151

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 152 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 211

QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAAACLLP 180
DB 212 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAAACLLP 271

QY 181 KLDELDEGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
DB 272 KLDELDEGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 331

QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCPEKPLEKSHCIAEVENDEMPA 300
DB 332 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCPEKPLEKSHCIAEVENDEMPA 391

QY 301 DLPSLAADFVSKDVCNKYAEAKDVLGMFLYFYARRHPDYSVVLNQLRLAKTYETTTLEKC 360
DB 392 DLPSLAADFVSKDVCNKYAEAKDVLGMFLYFYARRHPDYSVVLNQLRLAKTYETTTLEKC 451

QY 361 CAADDPHECYAKVDFEFPKPLVEEPQNLKQNCSELFQGLGEYKFNALLVRYTKKVPQVST 420
DB 452 CAADDPHECYAKVDFEFPKPLVEEPQNLKQNCSELFQGLGEYKFNALLVRYTKKVPQVST 511

QY 421 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 512 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 571

QY 481 LVNRRPCFSALEVDYETVYKPEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 540
DB 572 LVNRRPCFSALEVDYETVYKPEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 631

QY 541 KEQLKAVMDDFAAAFVEKCKCKADDKETCFABEGKGLVAASQAALGL 585
DB 632 KEQLKAVMDDFAAAFVEKCKCKADDKETCFABEGKGLVAASQAALGL 676

RESULT 57
US-10-153-064-129
; Sequence 129, Application US/10153064
; Publication No. US20020142814A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556

; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 129
; LENGTH: 676
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-129

Query Match 100.0%; Score 585; DB 13; Length 676;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRPKDGLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 92 DAHKEVAHRPKDGLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 151

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 152 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 211

QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAAACLLP 180
DB 212 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAAACLLP 271

QY 181 KLDELDEGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
DB 272 KLDELDEGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 331

QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCPEKPLEKSHCIAEVENDEMPA 300
DB 332 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCPEKPLEKSHCIAEVENDEMPA 391

QY 301 DLPSLAADFVSKDVCNKYAEAKDVLGMFLYFYARRHPDYSVVLNQLRLAKTYETTTLEKC 360
DB 392 DLPSLAADFVSKDVCNKYAEAKDVLGMFLYFYARRHPDYSVVLNQLRLAKTYETTTLEKC 451

QY 361 CAADDPHECYAKVDFEFPKPLVEEPQNLKQNCSELFQGLGEYKFNALLVRYTKKVPQVST 420
DB 452 CAADDPHECYAKVDFEFPKPLVEEPQNLKQNCSELFQGLGEYKFNALLVRYTKKVPQVST 511

QY 421 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 512 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 571

QY 481 LVNRRPCFSALEVDYETVYKPEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 540
DB 572 LVNRRPCFSALEVDYETVYKPEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 631

QY 541 KEQLKAVMDDFAAAFVEKCKCKADDKETCFABEGKGLVAASQAALGL 585
DB 632 KEQLKAVMDDFAAAFVEKCKCKADDKETCFABEGKGLVAASQAALGL 676

RESULT 58
US-10-153-604A-127
; Sequence 127, Application US/10153604A
; Publication No. US20030143191A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,604A
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 127
; LENGTH: 676

QY 541 KEQLKAVMDPFAAFVEKCKCKADKCTCFABEGKLVAAASQAALGL 585
 DB 542 KEQLKAVMDPFAAFVEKCKCKADKCTCFABEGKLVAAASQAALGL 586

RESULT 54
 US-10-425-000-34
 ; Sequence 34, Application US/10425000
 ; Publication No. US20040052777A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Nesbit, Mark
 ; APPLICANT: Cameron, Beatrice
 ; APPLICANT: Blanche, Francis
 ; TITLE OF INVENTION: Kingle Polypeptides and Methods for Using Them to Inhibit
 ; TITLE OF INVENTION: Angiogenesis
 ; FILE REFERENCE: ST01027-B
 ; CURRENT APPLICATION NUMBER: US/10/425,000
 ; CURRENT FILING DATE: 2003-04-29
 ; PRIOR APPLICATION NUMBER: 10/233,675
 ; PRIOR FILING DATE: 2002-09-04
 ; NUMBER OF SEQ ID NOS: 105
 ; SOFTWARE: Patent in version 3.2
 ; SEQ ID NO 34
 ; LENGTH: 674
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Human derived fusion protein
 ; US-10-425-000-34

Query Match 100.0%; Score 585; DB 12; Length 674;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQOCPPFEDHVKLVNEVTEFAKTCVADESAE 60
 DB 2 DAHSEVAHRFKDLGEENFKALVLIAPAYLQOCPPFEDHVKLVNEVTEFAKTCVADESAE 61
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 DB 62 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 121
 QY 121 DVMCTAFHNDNETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
 DB 122 DVMCTAFHNDNETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 181
 QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFKAFPAEVSCLVTDLT 240
 DB 182 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFKAFPAEVSCLVTDLT 241
 QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
 DB 242 VHTCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 301
 QY 301 DLPSLAADFVSKDVKCYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTLEK 360
 DB 302 DLPSLAADFVSKDVKCYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTLEK 361
 QY 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFGEYKFNALLVRYTKVPOVST 420
 DB 362 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFGEYKFNALLVRYTKVPOVST 421
 QY 421 PTLVEVSRLGKVGSKCKHPKAMPKCAEDYLSVNLQCLVHKEKTPVSDRVTKCCTES 480
 DB 422 PTLVEVSRLGKVGSKCKHPKAMPKCAEDYLSVNLQCLVHKEKTPVSDRVTKCCTES 481
 QY 481 LVNRRPCFSALEVDYVVPKEFNAETFTFHADICTLSEKERQIKKQATLVLVKHKPKAT 540
 DB 482 LVNRRPCFSALEVDYVVPKEFNAETFTFHADICTLSEKERQIKKQATLVLVKHKPKAT 541
 QY 541 KEQLKAVMDPFAAFVEKCKCKADKCTCFABEGKLVAAASQAALGL 585

Db 542 KEQLKAVMDPFAAFVEKCKCKADKCTCFABEGKLVAAASQAALGL 586

RESULT 55
 US-10-233-675A-14
 ; Sequence 14, Application US/10233675A
 ; Publication No. US20030228298A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Nesbit, Mark
 ; APPLICANT: Fong, Timothy
 ; APPLICANT: Brockstedt, Dirk
 ; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods f
 ; TITLE OF INVENTION: Throm To Inhibit Angiogenesis
 ; FILE REFERENCE: ST01027
 ; CURRENT APPLICATION NUMBER: US/10/233,675A
 ; CURRENT FILING DATE: 2002-09-04
 ; PRIOR APPLICATION NUMBER: 60/316,300
 ; PRIOR FILING DATE: 2001-09-04
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 14
 ; LENGTH: 674
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: fusion protein human abrogen
 ; US-10-233-675A-14

Query Match 100.0%; Score 585; DB 15; Length 674;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQOCPPFEDHVKLVNEVTEFAKTCVADESAE 60
 DB 2 DAHSEVAHRFKDLGEENFKALVLIAPAYLQOCPPFEDHVKLVNEVTEFAKTCVADESAE 61
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 DB 62 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 121
 QY 121 DVMCTAFHNDNETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
 DB 122 DVMCTAFHNDNETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 181
 QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFKAFPAEVSCLVTDLT 240
 DB 182 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFKAFPAEVSCLVTDLT 241
 QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
 DB 242 VHTCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 301
 QY 301 DLPSLAADFVSKDVKCYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTLEK 360
 DB 302 DLPSLAADFVSKDVKCYAEAKDVFLGMFLYFYARRHPDYSVVLLRLAKTYETTLEK 361
 QY 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFGEYKFNALLVRYTKVPOVST 420
 DB 362 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFGEYKFNALLVRYTKVPOVST 421
 QY 421 PTLVEVSRLGKVGSKCKHPKAMPKCAEDYLSVNLQCLVHKEKTPVSDRVTKCCTES 480
 DB 422 PTLVEVSRLGKVGSKCKHPKAMPKCAEDYLSVNLQCLVHKEKTPVSDRVTKCCTES 481
 QY 481 LVNRRPCFSALEVDYVVPKEFNAETFTFHADICTLSEKERQIKKQATLVLVKHKPKAT 540
 DB 482 LVNRRPCFSALEVDYVVPKEFNAETFTFHADICTLSEKERQIKKQATLVLVKHKPKAT 541
 QY 541 KEQLKAVMDPFAAFVEKCKCKADKCTCFABEGKLVAAASQAALGL 585
 DB 542 KEQLKAVMDPFAAFVEKCKCKADKCTCFABEGKLVAAASQAALGL 586

Db 508 PTLVEVSRNLKGVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 567
Qy 481 LVNRRPCFSALEVDVTVPKFNAETFTPHADICTLSEKEROIKKOTALVELVVKHKPKAT 540
Db 568 LVNRRPCFSALEVDVTVPKFNAETFTPHADICTLSEKEROIKKOTALVELVVKHKPKAT 627
Qy 541 KEQLKAVMDDPFAAFVCKCKADDKETCFABEGKKLVAAQAALGL 585
Db 628 KEQLKAVMDDPFAAFVCKCKADDKETCFABEGKKLVAAQAALGL 672

RESULT 52
US-10-233-675A-15
; Sequence 15, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; PRIOR FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 672
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion protein human abrogen
US-10-233-675A-15

Query Match 100.0%; Score 585; DB 15; Length 672;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKEVAHRRFKDLEENFKALVLIAPVLOQCPEDHVKLVNEVTFPAKTCVADESAAE 60
Db 86 DAHKEVAHRRFKDLEENFKALVLIAPVLOQCPEDHVKLVNEVTFPAKTCVADESAAE 147
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPPLRLVRPEV 120
Db 148 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPPLRLVRPEV 207
Qy 121 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
Db 208 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 267
Qy 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
Db 268 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 327
Qy 241 VHTCCCHGDLLECADRADLAKVICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 328 VHTCCCHGDLLECADRADLAKVICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 387
Qy 301 DLPSLAADPVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRLLAKTYETLEK 360
Db 388 DLPSLAADPVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRLLAKTYETLEK 447
Qy 361 CAADDPHECYAKVDFEPKPLVEEPQNLKONCELFQGLGEYKFNALLVRYTKVPQVST 420
Db 448 CAADDPHECYAKVDFEPKPLVEEPQNLKONCELFQGLGEYKFNALLVRYTKVPQVST 507
Qy 421 PTLVEVSRNLKGVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 508 PTLVEVSRNLKGVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 567

Qy 481 LVNRRPCFSALEVDVTVPKFNAETFTPHADICTLSEKEROIKKOTALVELVVKHKPKAT 540
Db 568 LVNRRPCFSALEVDVTVPKFNAETFTPHADICTLSEKEROIKKOTALVELVVKHKPKAT 627
Qy 541 KEQLKAVMDDPFAAFVCKCKADDKETCFABEGKKLVAAQAALGL 585
Db 628 KEQLKAVMDDPFAAFVCKCKADDKETCFABEGKKLVAAQAALGL 672

RESULT 53
US-10-424-999-14
; Sequence 14, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanchet, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-14

Query Match 100.0%; Score 585; DB 12; Length 674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKEVAHRRFKDLEENFKALVLIAPVLOQCPEDHVKLVNEVTFPAKTCVADESAAE 60
Db 2 DAHKEVAHRRFKDLEENFKALVLIAPVLOQCPEDHVKLVNEVTFPAKTCVADESAAE 61
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPPLRLVRPEV 120
Db 62 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPPLRLVRPEV 121
Qy 121 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
Db 122 DVNCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 181
Qy 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
Db 182 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 241
Qy 241 VHTCCCHGDLLECADRADLAKVICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 242 VHTCCCHGDLLECADRADLAKVICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 301
Qy 301 DLPSLAADPVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRLLAKTYETLEK 360
Db 302 DLPSLAADPVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRLLAKTYETLEK 361
Qy 361 CAADDPHECYAKVDFEPKPLVEEPQNLKONCELFQGLGEYKFNALLVRYTKVPQVST 420
Db 362 CAADDPHECYAKVDFEPKPLVEEPQNLKONCELFQGLGEYKFNALLVRYTKVPQVST 421
Qy 421 PTLVEVSRNLKGVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 422 PTLVEVSRNLKGVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 481
Qy 481 LVNRRPCFSALEVDVTVPKFNAETFTPHADICTLSEKEROIKKOTALVELVVKHKPKAT 540
Db 482 LVNRRPCFSALEVDVTVPKFNAETFTPHADICTLSEKEROIKKOTALVELVVKHKPKAT 541

QY 361 CAAADPHCYAKVDFEFPKLVPEEPQNLKQNCLEFQELGEYKFNQALLVRYTKKVPQVST 420
DB 432 CAAADPHCYAKVDFEFPKLVPEEPQNLKQNCLEFQELGEYKFNQALLVRYTKKVPQVST 491
QY 421 PTLVSVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQNLVHEKTPVSDRVTKCCTES 480
DB 492 PTLVSVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQNLVHEKTPVSDRVTKCCTES 551
QY 481 LVNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540
DB 552 LVNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 611
QY 541 KEQLKAVMDPFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
DB 612 KEQLKAVMDPFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 656

RESULT 50
US-10-424-999-15
; Sequence 15, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; PRIOR FILING DATE: 2003-04-29
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patent in version 3.2
; LENGTH: 672
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-15

Query Match 100.0%; Score 585; DB 12; Length 672;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFDLGEENFKALVLIAPAYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
DB 88 DAHSEVAHRFDLGEENFKALVLIAPAYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 147
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
DB 148 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 207
QY 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 208 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 267
QY 121 KLDELDEGKASAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTLDTLK 180
DB 208 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 267
QY 181 KLDELDEGKASAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTLDTLK 240
DB 268 KLDELDEGKASAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTLDTLK 327
QY 241 VHTCCHGDLLECCADDDRADLAKYICENQDSISSKLKECCCKPILKSHCIAEVENDEMPA 300
DB 328 VHTCCHGDLLECCADDDRADLAKYICENQDSISSKLKECCCKPILKSHCIAEVENDEMPA 387
QY 301 DLPSLAADFVSKDVKCNVAEAKDVLGMFLYEAARRHPDYSVLLLRKATYETTTLEKC 360
DB 388 DLPSLAADFVSKDVKCNVAEAKDVLGMFLYEAARRHPDYSVLLLRKATYETTTLEKC 447
QY 361 CAAADPHCYAKVDFEFPKLVPEEPQNLKQNCLEFQELGEYKFNQALLVRYTKKVPQVST 420

DB 448 CAAADPHCYAKVDFEFPKLVPEEPQNLKQNCLEFQELGEYKFNQALLVRYTKKVPQVST 507
QY 421 PTLVSVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQNLVHEKTPVSDRVTKCCTES 480
DB 508 PTLVSVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQNLVHEKTPVSDRVTKCCTES 567
QY 481 LVNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540
DB 568 LVNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 627
QY 541 KEQLKAVMDPFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
DB 628 KEQLKAVMDPFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 672

RESULT 51
US-10-425-000-35
; Sequence 35, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Kingle Polypeptides and Methods for Using Them to Inhibit
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 35
; LENGTH: 672
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-35

Query Match 100.0%; Score 585; DB 12; Length 672;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFDLGEENFKALVLIAPAYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
DB 88 DAHSEVAHRFDLGEENFKALVLIAPAYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 147
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
DB 148 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 207
QY 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 208 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 267
QY 181 KLDELDEGKASAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTLDTLK 240
DB 268 KLDELDEGKASAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTLDTLK 327
QY 241 VHTCCHGDLLECCADDDRADLAKYICENQDSISSKLKECCCKPILKSHCIAEVENDEMPA 300
DB 328 VHTCCHGDLLECCADDDRADLAKYICENQDSISSKLKECCCKPILKSHCIAEVENDEMPA 387
QY 301 DLPSLAADFVSKDVKCNVAEAKDVLGMFLYEAARRHPDYSVLLLRKATYETTTLEKC 360
DB 388 DLPSLAADFVSKDVKCNVAEAKDVLGMFLYEAARRHPDYSVLLLRKATYETTTLEKC 447
QY 361 CAAADPHCYAKVDFEFPKLVPEEPQNLKQNCLEFQELGEYKFNQALLVRYTKKVPQVST 420
DB 448 CAAADPHCYAKVDFEFPKLVPEEPQNLKQNCLEFQELGEYKFNQALLVRYTKKVPQVST 507
QY 421 PTLVSVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQNLVHEKTPVSDRVTKCCTES 480

QY 121 DVMCTAFHDNEETFLKKYLIEIARRHPFYFAPELLFFAKRYKAAFTSCCOADKAACLLP 180
DB 189 DVMCTAFHDNEETFLKKYLIEIARRHPFYFAPELLFFAKRYKAAFTSCCOADKAACLLP 248
QY 181 KLDELDEGKASSAKORLKASLOKFGBRAPKAWAVARLSORFPKABFAVSKLVTDLTK 240
DB 249 KLDELDEGKASSAKORLKASLOKFGBRAPKAWAVARLSORFPKABFAVSKLVTDLTK 308
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCCCKPPLLEKSHCIAEVENDEMPA 300
DB 309 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCCCKPPLLEKSHCIAEVENDEMPA 368
QY 301 DLPSLAADPVESKDVCKNYAEAKDVLGMFLFYIYARRHPDYSVVLRLRLAKTYETTLK 360
DB 369 DLPSLAADPVESKDVCKNYAEAKDVLGMFLFYIYARRHPDYSVVLRLRLAKTYETTLK 428
QY 361 CAAADPHECVAKVDFEFPKPLVEBPONLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
DB 429 CAAADPHECVAKVDFEFPKPLVEBPONLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 488
QY 421 PTLVEVSRNLGVSKCKCKHPKAMPKAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 489 PTLVEVSRNLGVSKCKCKHPKAMPKAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 548
QY 481 LVNRRPCFSALEVDVETVPKEFNAETFFHADICTLSEKEROIKKOTATLVELVKGKPKAT 540
DB 549 LVNRRPCFSALEVDVETVPKEFNAETFFHADICTLSEKEROIKKOTATLVELVKGKPKAT 608
QY 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
DB 609 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 653

RESULT 48
US-10-153-064-130
; Sequence 130, Application US/10153064
; Publication No. US20020142814A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 130
; LENGTH: 656
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-130

Query Match 100.0%; Score 585; DB 13; Length 656;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIATAFYAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60
DB 72 DAHKSEVAHRFKDLGEENFKALVLIATAFYAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 131
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPRVPEV 120
DB 132 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPRVPEV 191
QY 121 DVMCTAFHDNEETFLKKYLIEIARRHPFYFAPELLFFAKRYKAAFTSCCOADKAACLLP 180
DB 192 DVMCTAFHDNEETFLKKYLIEIARRHPFYFAPELLFFAKRYKAAFTSCCOADKAACLLP 251
QY 181 KLDELDEGKASSAKORLKASLOKFGBRAPKAWAVARLSORFPKABFAVSKLVTDLTK 240
DB 252 KLDELDEGKASSAKORLKASLOKFGBRAPKAWAVARLSORFPKABFAVSKLVTDLTK 311
QY 241 DVMCTAFHDNEETFLKKYLIEIARRHPFYFAPELLFFAKRYKAAFTSCCOADKAACLLP 300
DB 309 DVMCTAFHDNEETFLKKYLIEIARRHPFYFAPELLFFAKRYKAAFTSCCOADKAACLLP 368
QY 361 CAAADPHECVAKVDFEFPKPLVEBPONLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
DB 429 CAAADPHECVAKVDFEFPKPLVEBPONLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 488
QY 421 PTLVEVSRNLGVSKCKCKHPKAMPKAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 489 PTLVEVSRNLGVSKCKCKHPKAMPKAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 548
QY 481 LVNRRPCFSALEVDVETVPKEFNAETFFHADICTLSEKEROIKKOTATLVELVKGKPKAT 540
DB 549 LVNRRPCFSALEVDVETVPKEFNAETFFHADICTLSEKEROIKKOTATLVELVKGKPKAT 608
QY 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
DB 609 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 653

QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCCCKPPLLEKSHCIAEVENDEMPA 300
DB 312 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCCCKPPLLEKSHCIAEVENDEMPA 371
QY 301 DLPSLAADPVESKDVCKNYAEAKDVLGMFLFYIYARRHPDYSVVLRLRLAKTYETTLK 360
DB 372 DLPSLAADPVESKDVCKNYAEAKDVLGMFLFYIYARRHPDYSVVLRLRLAKTYETTLK 431
QY 361 CAAADPHECVAKVDFEFPKPLVEBPONLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
DB 432 CAAADPHECVAKVDFEFPKPLVEBPONLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 491
QY 421 PTLVEVSRNLGVSKCKCKHPKAMPKAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 492 PTLVEVSRNLGVSKCKCKHPKAMPKAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 551
QY 481 LVNRRPCFSALEVDVETVPKEFNAETFFHADICTLSEKEROIKKOTATLVELVKGKPKAT 540
DB 552 LVNRRPCFSALEVDVETVPKEFNAETFFHADICTLSEKEROIKKOTATLVELVKGKPKAT 611
QY 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
DB 612 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 656

RESULT 49
US-10-153-604A-130
; Sequence 130, Application US/10153604A
; Publication No. US20030143191A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,604A
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 130
; LENGTH: 656
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-604A-130

Query Match 100.0%; Score 585; DB 14; Length 656;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIATAFYAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60
DB 72 DAHKSEVAHRFKDLGEENFKALVLIATAFYAQYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 131
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPRVPEV 120
DB 132 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPRVPEV 191
QY 121 DVMCTAFHDNEETFLKKYLIEIARRHPFYFAPELLFFAKRYKAAFTSCCOADKAACLLP 180
DB 192 DVMCTAFHDNEETFLKKYLIEIARRHPFYFAPELLFFAKRYKAAFTSCCOADKAACLLP 251
QY 181 KLDELDEGKASSAKORLKASLOKFGBRAPKAWAVARLSORFPKABFAVSKLVTDLTK 240
DB 252 KLDELDEGKASSAKORLKASLOKFGBRAPKAWAVARLSORFPKABFAVSKLVTDLTK 311
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCCCKPPLLEKSHCIAEVENDEMPA 300
DB 312 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCCCKPPLLEKSHCIAEVENDEMPA 371
QY 301 DLPSLAADPVESKDVCKNYAEAKDVLGMFLFYIYARRHPDYSVVLRLRLAKTYETTLK 360
DB 372 DLPSLAADPVESKDVCKNYAEAKDVLGMFLFYIYARRHPDYSVVLRLRLAKTYETTLK 431

```

; LENGTH: 652
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-604A-132

Query Match      100.0%; Score 585; DB 14; Length 652;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDILGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 69 DAHSEVAHRFKDILGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 128

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLRLVRPEV 120
DB 129 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLRLVRPEV 188

QY 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKYKAAFTCCQAAADKAACLLP 180
DB 189 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKYKAAFTCCQAAADKAACLLP 248

QY 181 KLDELRLDEGKASSAKORLKCSLOKFGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 240
DB 249 KLDELRLDEGKASSAKORLKCSLOKFGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 308

QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDMPA 300
DB 309 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDMPA 368

QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLILRLAKTYETTLKX 360
DB 369 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLILRLAKTYETTLKX 428

QY 361 CAADAPHECYAKVDFEFPKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKVPQVST 420
DB 429 CAADAPHECYAKVDFEFPKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKVPQVST 488

QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCCTES 480
DB 489 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCCTES 548

QY 481 LVNRRPCFSALEVDETVYVPKEFNAETFFHADICTLSEKERQIKQTALVELVKHHPKAT 540
DB 549 LVNRRPCFSALEVDETVYVPKEFNAETFFHADICTLSEKERQIKQTALVELVKHHPKAT 608

QY 541 KEQLKAVMDDFAAFEVKCKKADDKETCFABEGKKLVAASQAALGL 585
DB 609 KEQLKAVMDDFAAFEVKCKKADDKETCFABEGKKLVAASQAALGL 653

RESULT 47
US-10-153-604A-131
; Sequence 131, Application US/10153604A
; Publication No. US20030143191A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,604A
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 131
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-604A-131

Query Match      100.0%; Score 585; DB 14; Length 653;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDILGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 69 DAHSEVAHRFKDILGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 128

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLRLVRPEV 120
DB 129 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLRLVRPEV 188

QY 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKYKAAFTCCQAAADKAACLLP 180
DB 189 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKYKAAFTCCQAAADKAACLLP 248

QY 181 KLDELRLDEGKASSAKORLKCSLOKFGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 240
DB 249 KLDELRLDEGKASSAKORLKCSLOKFGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 308

QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDMPA 300
DB 309 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDMPA 368

QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLILRLAKTYETTLKX 360
DB 369 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLILRLAKTYETTLKX 428

QY 361 CAADAPHECYAKVDFEFPKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKVPQVST 420
DB 429 CAADAPHECYAKVDFEFPKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKVPQVST 488

QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCCTES 480
DB 489 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCCTES 548

QY 481 LVNRRPCFSALEVDETVYVPKEFNAETFFHADICTLSEKERQIKQTALVELVKHHPKAT 540
DB 549 LVNRRPCFSALEVDETVYVPKEFNAETFFHADICTLSEKERQIKQTALVELVKHHPKAT 608

QY 541 KEQLKAVMDDFAAFEVKCKKADDKETCFABEGKKLVAASQAALGL 585
DB 609 KEQLKAVMDDFAAFEVKCKKADDKETCFABEGKKLVAASQAALGL 653

RESULT 46
US-10-153-064-131
; Sequence 131, Application US/10153064
; Publication No. US20020142814A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 131
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-131

Query Match      100.0%; Score 585; DB 13; Length 653;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
RESULT 43
US-10-153-604A-133
; Sequence 133, Application US/10153604A
; Publication No. US20030143191A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 133
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-604A-133

Query Match      100.0%; Score 585; DB 14; Length 651;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDILGEENFKALVLIAPFAQYLOQCPEDEHVKLVNEVTEFAKTCVADESAAE 60
Db 67 DAHSEVAHRFKDILGEENFKALVLIAPFAQYLOQCPEDEHVKLVNEVTEFAKTCVADESAAE 126

QY 51 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKOBERNECEFLQHKDNDPNLPRLVRREV 120
Db 127 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKOBERNECEFLQHKDNDPNLPRLVRREV 186

QY 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 187 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 246

QY 181 KLDELRDGKASSAKORLKASLOKGERAFKAWAVARLSQFPKAEVSKLVTDLTK 240
Db 247 KLDELRDGKASSAKORLKASLOKGERAFKAWAVARLSQFPKAEVSKLVTDLTK 306

QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKKECEKPLLEKSHCIAEVENDEMPA 300
Db 307 VHTCCCHGDLLECADRADLAKYICENQDSISSKKECEKPLLEKSHCIAEVENDEMPA 366

QY 301 DPLSLAADFVSKDVCNKYAEADKVLGMFLYEAARRHPDYSVLLLRLLAKTYETTLKCC 360
Db 367 DPLSLAADFVSKDVCNKYAEADKVLGMFLYEAARRHPDYSVLLLRLLAKTYETTLKCC 426

QY 361 CAADDPHECYAKVDFPKPLVEBPQNLIKQNCFLFQGLGEYKFNALLVRYTKVPQVST 420
Db 427 CAADDPHECYAKVDFPKPLVEBPQNLIKQNCFLFQGLGEYKFNALLVRYTKVPQVST 486

QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHKTVPVSDRVTKCCTES 480
Db 487 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHKTVPVSDRVTKCCTES 546

QY 481 LVNRRPCFSALVEVDYVPKFEFNAETFTFHADICTLSEKERQIKQTALVELVGHKPKAT 540
Db 547 LVNRRPCFSALVEVDYVPKFEFNAETFTFHADICTLSEKERQIKQTALVELVGHKPKAT 606

QY 541 KEOLKAVMDDFAAFVEKCKCADDKETCFABEGKKLVAASQAALGL 585
Db 607 KEOLKAVMDDFAAFVEKCKCADDKETCFABEGKKLVAASQAALGL 651

RESULT 44
US-10-153-604A-132
; Sequence 132, Application US/10153604A
; Publication No. US20020142814A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 132
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-604A-132

Query Match      100.0%; Score 585; DB 13; Length 652;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDILGEENFKALVLIAPFAQYLOQCPEDEHVKLVNEVTEFAKTCVADESAAE 60
Db 68 DAHSEVAHRFKDILGEENFKALVLIAPFAQYLOQCPEDEHVKLVNEVTEFAKTCVADESAAE 127

QY 51 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKOBERNECEFLQHKDNDPNLPRLVRREV 120
Db 128 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKOBERNECEFLQHKDNDPNLPRLVRREV 187

QY 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 188 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 247

QY 181 KLDELRDGKASSAKORLKASLOKGERAFKAWAVARLSQFPKAEVSKLVTDLTK 240
Db 248 KLDELRDGKASSAKORLKASLOKGERAFKAWAVARLSQFPKAEVSKLVTDLTK 307

QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKKECEKPLLEKSHCIAEVENDEMPA 300
Db 308 VHTCCCHGDLLECADRADLAKYICENQDSISSKKECEKPLLEKSHCIAEVENDEMPA 367

QY 301 DPLSLAADFVSKDVCNKYAEADKVLGMFLYEAARRHPDYSVLLLRLLAKTYETTLKCC 360
Db 368 DPLSLAADFVSKDVCNKYAEADKVLGMFLYEAARRHPDYSVLLLRLLAKTYETTLKCC 427

QY 361 CAADDPHECYAKVDFPKPLVEBPQNLIKQNCFLFQGLGEYKFNALLVRYTKVPQVST 420
Db 428 CAADDPHECYAKVDFPKPLVEBPQNLIKQNCFLFQGLGEYKFNALLVRYTKVPQVST 487

QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHKTVPVSDRVTKCCTES 480
Db 488 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHKTVPVSDRVTKCCTES 547

QY 481 LVNRRPCFSALVEVDYVPKFEFNAETFTFHADICTLSEKERQIKQTALVELVGHKPKAT 540
Db 548 LVNRRPCFSALVEVDYVPKFEFNAETFTFHADICTLSEKERQIKQTALVELVGHKPKAT 607

QY 541 KEOLKAVMDDFAAFVEKCKCADDKETCFABEGKKLVAASQAALGL 585
Db 608 KEOLKAVMDDFAAFVEKCKCADDKETCFABEGKKLVAASQAALGL 652

RESULT 45
US-10-153-604A-132
; Sequence 132, Application US/10153604A
; Publication No. US20030143191A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 132
```

Db 416 CAADPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQELGEYKFNQALLVRYTKKVPQVST 475
Qy 421 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDYLSVNLQLCVLHEKTPVSDRVTKCCTES 480
Db 476 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDYLSVNLQLCVLHEKTPVSDRVTKCCTES 535
Qy 481 LVNRRPCFSALEVDETVVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540
Db 536 LVNRRPCFSALEVDETVVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 595
Qy 541 KEQLKAVMDDFAAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 585
Db 596 KEQLKAVMDDFAAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 640

RESULT 41
US-10-433-108-17
; Sequence 17, Application US/10433108
; Publication No. US20040053370A1
; GENERAL INFORMATION:
; APPLICANT: Eli Lilly and Company
; TITLE OF INVENTION: GUP-1 FUSION PROTEINS
; FILE REFERENCE: X-13991
; CURRENT APPLICATION NUMBER: US/10/433,108
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: US 60/251,954
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 17
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-433-108-17

Query Match 100.0%; Score 585; DB 12; Length 640;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVLIAPQYLOQCFEDHVKLVNEVTEPAKTCVADESAE 60
Db 56 DAHKSEVAHRFKDLGEENFKALVLIAPQYLOQCFEDHVKLVNEVTEPAKTCVADESAE 115
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 116 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 175
Qy 121 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Db 176 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 235
Qy 181 KLDELDRDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 240
Db 236 KLDELDRDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 295
Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPILLESKSHCIAEVNDEMFA 300
Db 296 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPILLESKSHCIAEVNDEMFA 355
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEVARRHPDYSVVLRLRLAKTYETTTLEKC 360
Db 356 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEVARRHPDYSVVLRLRLAKTYETTTLEKC 415
Qy 361 CAADPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQELGEYKFNQALLVRYTKKVPQVST 420
Db 416 CAADPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQELGEYKFNQALLVRYTKKVPQVST 475
Qy 421 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDYLSVNLQLCVLHEKTPVSDRVTKCCTES 480
Db 476 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDYLSVNLQLCVLHEKTPVSDRVTKCCTES 535

Qy 481 LVNRRPCFSALEVDETVVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540
Db 536 LVNRRPCFSALEVDETVVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 595
Qy 541 KEQLKAVMDDFAAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 585
Db 596 KEQLKAVMDDFAAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 640

RESULT 42
US-10-153-064-133
; Sequence 133, Application US/10153064
; Publication No. US20020142814A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 133
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-133

Query Match 100.0%; Score 585; DB 13; Length 651;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVLIAPQYLOQCFEDHVKLVNEVTEPAKTCVADESAE 60
Db 67 DAHKSEVAHRFKDLGEENFKALVLIAPQYLOQCFEDHVKLVNEVTEPAKTCVADESAE 126
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 127 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 186
Qy 121 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Db 187 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 246
Qy 181 KLDELDRDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 240
Db 247 KLDELDRDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 306
Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPILLESKSHCIAEVNDEMFA 300
Db 307 VHTCCCHGDLLECADRADLAKYICENQDSISSKLECCCKPILLESKSHCIAEVNDEMFA 366
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEVARRHPDYSVVLRLRLAKTYETTTLEKC 360
Db 367 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYEVARRHPDYSVVLRLRLAKTYETTTLEKC 426
Qy 361 CAADPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQELGEYKFNQALLVRYTKKVPQVST 420
Db 427 CAADPHECYAKVDFEFPKPLVEEPQNLIKQNCLEFQELGEYKFNQALLVRYTKKVPQVST 486
Qy 421 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDYLSVNLQLCVLHEKTPVSDRVTKCCTES 480
Db 487 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDYLSVNLQLCVLHEKTPVSDRVTKCCTES 546
Qy 481 LVNRRPCFSALEVDETVVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540
Db 547 LVNRRPCFSALEVDETVVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 606
Qy 541 KEQLKAVMDDFAAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 585
Db 607 KEQLKAVMDDFAAAFVEKCKKADDDKTCFAEEGKKLVAASQAALGL 651

Db 220 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRFKABFAEVSCLVTLDTK 279
Qy 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECKEPLLEKSHCIAEVENDEMPA 300
Db 280 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECKEPLLEKSHCIAEVENDEMPA 339
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYVEYARRHPDYSVVLLRLAKTYETTLK 360
Db 340 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYVEYARRHPDYSVVLLRLAKTYETTLK 399
Qy 361 CAADPHCEYAKVDFBPKPLVEBPONLIKONCELPQOLGEYKFNALLVRYTKVPQVST 420
Db 400 CAADPHCEYAKVDFBPKPLVEBPONLIKONCELPQOLGEYKFNALLVRYTKVPQVST 459
Qy 421 PTLVEVSRLNLGVSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 460 PTLVEVSRLNLGVSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 519
Qy 481 LVNRRPCFSALEVDYVYPKEFNAETFFHADICTLSEKEROIKKQOTALVELVGHKPKAT 540
Db 520 LVNRRPCFSALEVDYVYPKEFNAETFFHADICTLSEKEROIKKQOTALVELVGHKPKAT 579
Qy 541 KEOLKAVMDDFAAFVEKCKCKADDDKCTCFABEGKCLVAASQAALGL 585
Db 580 KEOLKAVMDDFAAFVEKCKCKADDDKCTCFABEGKCLVAASQAALGL 624
RESULT 39
US-10-433-108-14
; Sequence 14, Application US/10433108
; Publication No. US20040053370A1
; GENERAL INFORMATION:
; APPLICANT: Eli Lilly and Company
; TITLE OF INVENTION: GLP-1 FUSION PROTEINS
; FILE REFERENCE: X-13991
; CURRENT APPLICATION NUMBER: US/10/433,108
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: US 60/251,954
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 14
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-433-108-14
Query Match 100.0%; Score 585; DB 12; Length 631;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLQOCPPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 47 DAHKSEVAHRFKDLGEENFKALVLIAPAYLQOCPPFEDHVKLVNEVTEFAKTCVADESAAE 106
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 120
Db 107 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 166
Qy 121 DVNCTAFHNDNEETFLKYLVEIARRHPYFYAPPELLFPKRYKAAFTCCQADKAAACLLP 180
Db 167 DVNCTAFHNDNEETFLKYLVEIARRHPYFYAPPELLFPKRYKAAFTCCQADKAAACLLP 226
Qy 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRFKABFAEVSCLVTLDTK 240
Db 227 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRFKABFAEVSCLVTLDTK 286
Qy 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECKEPLLEKSHCIAEVENDEMPA 300
Db 287 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECKEPLLEKSHCIAEVENDEMPA 346

Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYVEYARRHPDYSVVLLRLAKTYETTLK 360
Db 347 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYVEYARRHPDYSVVLLRLAKTYETTLK 406
Qy 361 CAADPHCEYAKVDFBPKPLVEBPONLIKONCELPQOLGEYKFNALLVRYTKVPQVST 420
Db 407 CAADPHCEYAKVDFBPKPLVEBPONLIKONCELPQOLGEYKFNALLVRYTKVPQVST 466
Qy 421 PTLVEVSRLNLGVSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 467 PTLVEVSRLNLGVSKCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 526
Qy 481 LVNRRPCFSALEVDYVYPKEFNAETFFHADICTLSEKEROIKKQOTALVELVGHKPKAT 540
Db 527 LVNRRPCFSALEVDYVYPKEFNAETFFHADICTLSEKEROIKKQOTALVELVGHKPKAT 586
Qy 541 KEOLKAVMDDFAAFVEKCKCKADDDKCTCFABEGKCLVAASQAALGL 585
Db 587 KEOLKAVMDDFAAFVEKCKCKADDDKCTCFABEGKCLVAASQAALGL 631
RESULT 40
US-10-433-108-15
; Sequence 15, Application US/10433108
; Publication No. US20040053370A1
; GENERAL INFORMATION:
; APPLICANT: Eli Lilly and Company
; TITLE OF INVENTION: GLP-1 FUSION PROTEINS
; FILE REFERENCE: X-13991
; CURRENT APPLICATION NUMBER: US/10/433,108
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: US 60/251,954
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 15
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-433-108-15
Query Match 100.0%; Score 585; DB 12; Length 640;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLQOCPPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 56 DAHKSEVAHRFKDLGEENFKALVLIAPAYLQOCPPFEDHVKLVNEVTEFAKTCVADESAAE 115
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 120
Db 116 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 175
Qy 121 DVNCTAFHNDNEETFLKYLVEIARRHPYFYAPPELLFPKRYKAAFTCCQADKAAACLLP 180
Db 176 DVNCTAFHNDNEETFLKYLVEIARRHPYFYAPPELLFPKRYKAAFTCCQADKAAACLLP 235
Qy 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRFKABFAEVSCLVTLDTK 240
Db 236 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRFKABFAEVSCLVTLDTK 295
Qy 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECKEPLLEKSHCIAEVENDEMPA 300
Db 296 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECKEPLLEKSHCIAEVENDEMPA 355
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYVEYARRHPDYSVVLLRLAKTYETTLK 360
Db 356 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYVEYARRHPDYSVVLLRLAKTYETTLK 415
Qy 361 CAADPHCEYAKVDFBPKPLVEBPONLIKONCELPQOLGEYKFNALLVRYTKVPQVST 420

Db 25 DAHSEVAHRFKDLGEEFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADBSAE 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLRLVRPEV 144
Qy 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTBCCQAADKAACLLP 180
Db 145 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTBCCQAADKAACLLP 204
Qy 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 240
Db 205 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 264
Qy 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCPEKPLEKSHCIAEVENDEMPA 300
Db 265 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCPEKPLEKSHCIAEVENDEMPA 324
Qy 301 DLPSLAADFVESKDVCKNYABAKDVLGMLFYEYARRHPDYSVLLLRKAKTYETTLTK 360
Db 325 DLPSLAADFVESKDVCKNYABAKDVLGMLFYEYARRHPDYSVLLLRKAKTYETTLTK 384
Qy 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFEQLGKFKQNALLVRYTKVPOVST 420
Db 385 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFEQLGKFKQNALLVRYTKVPOVST 444
Qy 421 PTLVSVRNLGKVGSKCKCHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVSVRNLGKVGSKCKCHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504
Qy 481 LVNRRPCFSALEVDYVYKPEFNAETTFHADICTLSEKEROIKKOTALVELVKHKPKAT 540
Db 505 LVNRRPCFSALEVDYVYKPEFNAETTFHADICTLSEKEROIKKOTALVELVKHKPKAT 564
Qy 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKGLVAASQAALGL 585
Db 565 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKGLVAASQAALGL 609

RESULT 37
US-10-433-108-13
; Sequence 13, Application US/10433108
; Publication No. US20040053370A1
; GENERAL INFORMATION:
; APPLICANT: Eli Lilly and Company
; TITLE OF INVENTION: GLP-1 FUSION PROTEINS
; FILE REFERENCE: X-13991
; CURRENT APPLICATION NUMBER: US/10/433,108
; PRIOR FILING DATE: 2003-05-29
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-433-108-13

Query Match 100.0%; Score 585; DB 12; Length 616;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DAHSEVAHRFKDLGEEFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADBSAE 60
Db 32 DAHSEVAHRFKDLGEEFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADBSAE 91
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLRLVRPEV 120
Db 92 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLRLVRPEV 151

Qy 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTBCCQAADKAACLLP 180
Db 152 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTBCCQAADKAACLLP 211
Qy 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 240
Db 212 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 271
Qy 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCPEKPLEKSHCIAEVENDEMPA 300
Db 272 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCPEKPLEKSHCIAEVENDEMPA 331
Qy 301 DLPSLAADFVESKDVCKNYABAKDVLGMLFYEYARRHPDYSVLLLRKAKTYETTLTK 360
Db 332 DLPSLAADFVESKDVCKNYABAKDVLGMLFYEYARRHPDYSVLLLRKAKTYETTLTK 391
Qy 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFEQLGKFKQNALLVRYTKVPOVST 420
Db 392 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFEQLGKFKQNALLVRYTKVPOVST 451
Qy 421 PTLVSVRNLGKVGSKCKCHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 452 PTLVSVRNLGKVGSKCKCHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 511
Qy 481 LVNRRPCFSALEVDYVYKPEFNAETTFHADICTLSEKEROIKKOTALVELVKHKPKAT 540
Db 512 LVNRRPCFSALEVDYVYKPEFNAETTFHADICTLSEKEROIKKOTALVELVKHKPKAT 571
Qy 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKGLVAASQAALGL 585
Db 572 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKGLVAASQAALGL 616

RESULT 38
US-10-433-108-16
; Sequence 16, Application US/10433108
; Publication No. US20040053370A1
; GENERAL INFORMATION:
; APPLICANT: Eli Lilly and Company
; TITLE OF INVENTION: GLP-1 FUSION PROTEINS
; FILE REFERENCE: X-13991
; CURRENT APPLICATION NUMBER: US/10/433,108
; PRIOR FILING DATE: 2003-05-29
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-433-108-16

Query Match 100.0%; Score 585; DB 12; Length 624;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DAHSEVAHRFKDLGEEFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADBSAE 60
Db 40 DAHSEVAHRFKDLGEEFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADBSAE 99
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLRLVRPEV 120
Db 100 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLRLVRPEV 159
Qy 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTBCCQAADKAACLLP 180
Db 160 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPELLFFAKRYKAAFTBCCQAADKAACLLP 219
Qy 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 240

Db 145 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLOKQGERAFKAWAVARLSORFFKABAEVSKLVTDLTK 240
Db 205 KLDELDEGKASSAKQRLKASLOKQGERAFKAWAVARLSORFFKABAEVSKLVTDLTK 264
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKKECEKPLLEKSHCIAEVNDMPA 300
Db 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKKECEKPLLEKSHCIAEVNDMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLVEYARRHPDYVSVLLNQLCVLHEKTPVSDRVTKCCTES 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVLGMFLVEYARRHPDYVSVLLNQLCVLHEKTPVSDRVTKCCTES 384
QY 361 CAADADHECYAKVDFEPLVVEPQNLIKQNCLEFELQGEYKFQNALVRYTKVPQVST 420
Db 385 CAADADHECYAKVDFEPLVVEPQNLIKQNCLEFELQGEYKFQNALVRYTKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKEPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKEPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDETVYVPKEFNAETFTFHADICTLSEKEROIKKQOTALVELVHKPKAT 540
Db 505 LVNRRPCFSALEVDETVYVPKEFNAETFTFHADICTLSEKEROIKKQOTALVELVHKPKAT 564
QY 541 KEOLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585
Db 565 KEOLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 609

RESULT 34

US-10-237-866-2
; Sequence 2, Application US/10237866
; Publication No. US20030036171A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittou, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,866
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619

REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-237-866-2

Query Match 100.0%; Score 585; DB 14; Length 610;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHAFKDLGEENFKALVLIAPVLOQCPEDHVKLVNEVTEFAKTCVADESAE 60
Db 25 DAHSEVAHAFKDLGEENFKALVLIAPVLOQCPEDHVKLVNEVTEFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBERNECFLOHKDDPNLPRLVRREV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBERNECFLOHKDDPNLPRLVRREV 144
QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 145 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLOKQGERAFKAWAVARLSORFFKABAEVSKLVTDLTK 240
Db 205 KLDELDEGKASSAKQRLKASLOKQGERAFKAWAVARLSORFFKABAEVSKLVTDLTK 264
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKKECEKPLLEKSHCIAEVNDMPA 300
Db 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKKECEKPLLEKSHCIAEVNDMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLVEYARRHPDYVSVLLNQLCVLHEKTPVSDRVTKCCTES 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVLGMFLVEYARRHPDYVSVLLNQLCVLHEKTPVSDRVTKCCTES 384
QY 361 CAADADHECYAKVDFEPLVVEPQNLIKQNCLEFELQGEYKFQNALVRYTKVPQVST 420
Db 385 CAADADHECYAKVDFEPLVVEPQNLIKQNCLEFELQGEYKFQNALVRYTKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKEPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKEPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDETVYVPKEFNAETFTFHADICTLSEKEROIKKQOTALVELVHKPKAT 540
Db 505 LVNRRPCFSALEVDETVYVPKEFNAETFTFHADICTLSEKEROIKKQOTALVELVHKPKAT 564
QY 541 KEOLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585
Db 565 KEOLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 609

RESULT 35

US-10-237-871-2
; Sequence 2, Application US/10237871
; Publication No. US20030036172A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittou, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:

```
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/10/237,667
  FILING DATE: 10-Sep-2002
  CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US/08/797,689
  FILING DATE: 31-JAN-1997
  APPLICATION NUMBER: US 08/256,927
  FILING DATE: 28-JUL-1994
  APPLICATION NUMBER: FR 92/01064
  FILING DATE: 31-JAN-1992
  APPLICATION NUMBER: PCT/FR93/00085
  FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
  NAME: Smith Ph.D., Julie K.
  REGISTRATION NUMBER: P-38,619
  REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (610) 454-3839
  TELEFAX: (610) 454-3808
SEQUENCE CHARACTERISTICS:
  LENGTH: 610 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
  MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-237-667-2

Query Match      100.0%; Score 585; DB 14; Length 610;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAFRFKDLGSENFKALVLIAPAYLQOCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHSEVAFRFKDLGSENFKALVLIAPAYLQOCPFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDNNPLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDNNPLRLVRPEV 144
QY 121 DVNCTAFPHNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 145 DVNCTAFPHNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLQKGFGERAFKAMAVARLSQRFPAKAEPAEVSCLVTLTK 240
DB 205 KLDELDEGKASSAKQRLKASLQKGFGERAFKAMAVARLSQRFPAKAEPAEVSCLVTLTK 264
QY 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
DB 265 VHTECCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCNVAEAKDVFGLNFLYEYARRHPDYSVVLLLRKLTAKTYETTLK 360
DB 325 DLPSLAADFVESKDVCNVAEAKDVFGLNFLYEYARRHPDYSVVLLLRKLTAKTYETTLK 384
QY 361 CAADPHCEYAKVDFEFKPLVEBPQNLIKONCELFQELGEYKFNQALLVRYTKVPQVST 420
DB 385 CAADPHCEYAKVDFEFKPLVEBPQNLIKONCELFQELGEYKFNQALLVRYTKVPQVST 444
QY 421 PTLVEVSRLNGKVGSKCKGHPKAKMPCAEEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRLNGKVGSKCKGHPKAKMPCAEEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 504
QY 481 LVNRRPFCFSALEVDETVVPKFEFNAETFTTHADICTLSEKEROIKKQTALVELVKHKPKAT 540
DB 505 LVNRRPFCFSALEVDETVVPKFEFNAETFTTHADICTLSEKEROIKKQTALVELVKHKPKAT 564
QY 541 KEQLKAWMDFAAFVEKCKCKADDDKTCFAEEGKKLVAAASQAALGL 585

COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/10/237,708
  FILING DATE: 10-Sep-2002
  CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US/08/797,689
  FILING DATE: 31-JAN-1997
  APPLICATION NUMBER: US 08/256,927
  FILING DATE: 28-JUL-1994
  APPLICATION NUMBER: FR 92/01064
  FILING DATE: 31-JAN-1992
  APPLICATION NUMBER: PCT/FR93/00085
  FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
  NAME: Smith Ph.D., Julie K.
  REGISTRATION NUMBER: P-38,619
  REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (610) 454-3839
  TELEFAX: (610) 454-3808
SEQUENCE CHARACTERISTICS:
  LENGTH: 610 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
  MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-237-708-2

Query Match      100.0%; Score 585; DB 14; Length 610;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAFRFKDLGSENFKALVLIAPAYLQOCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHSEVAFRFKDLGSENFKALVLIAPAYLQOCPFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDNNPLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDNNPLRLVRPEV 144
QY 121 DVNCTAFPHNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 145 DVNCTAFPHNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLQKGFGERAFKAMAVARLSQRFPAKAEPAEVSCLVTLTK 240
DB 205 KLDELDEGKASSAKQRLKASLQKGFGERAFKAMAVARLSQRFPAKAEPAEVSCLVTLTK 264
QY 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
DB 265 VHTECCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCNVAEAKDVFGLNFLYEYARRHPDYSVVLLLRKLTAKTYETTLK 360
DB 325 DLPSLAADFVESKDVCNVAEAKDVFGLNFLYEYARRHPDYSVVLLLRKLTAKTYETTLK 384
QY 361 CAADPHCEYAKVDFEFKPLVEBPQNLIKONCELFQELGEYKFNQALLVRYTKVPQVST 420
DB 385 CAADPHCEYAKVDFEFKPLVEBPQNLIKONCELFQELGEYKFNQALLVRYTKVPQVST 444
QY 421 PTLVEVSRLNGKVGSKCKGHPKAKMPCAEEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRLNGKVGSKCKGHPKAKMPCAEEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 504
QY 481 LVNRRPFCFSALEVDETVVPKFEFNAETFTTHADICTLSEKEROIKKQTALVELVKHKPKAT 540
DB 505 LVNRRPFCFSALEVDETVVPKFEFNAETFTTHADICTLSEKEROIKKQTALVELVKHKPKAT 564
QY 541 KEQLKAWMDFAAFVEKCKCKADDDKTCFAEEGKKLVAAASQAALGL 585

COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/10/237,708
  FILING DATE: 10-Sep-2002
  CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US/08/797,689
  FILING DATE: 31-JAN-1997
  APPLICATION NUMBER: US 08/256,927
  FILING DATE: 28-JUL-1994
  APPLICATION NUMBER: FR 92/01064
  FILING DATE: 31-JAN-1992
  APPLICATION NUMBER: PCT/FR93/00085
  FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
  NAME: Smith Ph.D., Julie K.
  REGISTRATION NUMBER: P-38,619
  REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (610) 454-3839
  TELEFAX: (610) 454-3808
SEQUENCE CHARACTERISTICS:
  LENGTH: 610 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
  MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-237-708-2

Query Match      100.0%; Score 585; DB 14; Length 610;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAFRFKDLGSENFKALVLIAPAYLQOCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHSEVAFRFKDLGSENFKALVLIAPAYLQOCPFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDNNPLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDNNPLRLVRPEV 144
QY 121 DVNCTAFPHNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 145 DVNCTAFPHNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLQKGFGERAFKAMAVARLSQRFPAKAEPAEVSCLVTLTK 240
DB 205 KLDELDEGKASSAKQRLKASLQKGFGERAFKAMAVARLSQRFPAKAEPAEVSCLVTLTK 264
QY 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
DB 265 VHTECCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCNVAEAKDVFGLNFLYEYARRHPDYSVVLLLRKLTAKTYETTLK 360
DB 325 DLPSLAADFVESKDVCNVAEAKDVFGLNFLYEYARRHPDYSVVLLLRKLTAKTYETTLK 384
QY 361 CAADPHCEYAKVDFEFKPLVEBPQNLIKONCELFQELGEYKFNQALLVRYTKVPQVST 420
DB 385 CAADPHCEYAKVDFEFKPLVEBPQNLIKONCELFQELGEYKFNQALLVRYTKVPQVST 444
QY 421 PTLVEVSRLNGKVGSKCKGHPKAKMPCAEEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRLNGKVGSKCKGHPKAKMPCAEEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 504
QY 481 LVNRRPFCFSALEVDETVVPKFEFNAETFTTHADICTLSEKEROIKKQTALVELVKHKPKAT 540
DB 505 LVNRRPFCFSALEVDETVVPKFEFNAETFTTHADICTLSEKEROIKKQTALVELVKHKPKAT 564
QY 541 KEQLKAWMDFAAFVEKCKCKADDDKTCFAEEGKKLVAAASQAALGL 585

COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/10/237,708
  FILING DATE: 10-Sep-2002
  CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US/08/797,689
  FILING DATE: 31-JAN-1997
  APPLICATION NUMBER: US 08/256,927
  FILING DATE: 28-JUL-1994
  APPLICATION NUMBER: FR 92/01064
  FILING DATE: 31-JAN-1992
  APPLICATION NUMBER: PCT/FR93/00085
  FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
  NAME: Smith Ph.D., Julie K.
  REGISTRATION NUMBER: P-38,619
  REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (610) 454-3839
  TELEFAX: (610) 454-3808
SEQUENCE CHARACTERISTICS:
  LENGTH: 610 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
  MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-237-708-2

Query Match      100.0%; Score 585; DB 14; Length 610;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAFRFKDLGSENFKALVLIAPAYLQOCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHSEVAFRFKDLGSENFKALVLIAPAYLQOCPFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDNNPLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDNNPLRLVRPEV 144
QY 121 DVNCTAFPHNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 145 DVNCTAFPHNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLQKGFGERAFKAMAVARLSQRFPAKAEPAEVSCLVTLTK 240
DB 205 KLDELDEGKASSAKQRLKASLQKGFGERAFKAMAVARLSQRFPAKAEPAEVSCLVTLTK 264
QY 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
DB 265 VHTECCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCNVAEAKDVFGLNFLYEYARRHPDYSVVLLLRKLTAKTYETTLK 360
DB 325 DLPSLAADFVESKDVCNVAEAKDVFGLNFLYEYARRHPDYSVVLLLRKLTAKTYETTLK 384
QY 361 CAADPHCEYAKVDFEFKPLVEBPQNLIKONCELFQELGEYKFNQALLVRYTKVPQVST 420
DB 385 CAADPHCEYAKVDFEFKPLVEBPQNLIKONCELFQELGEYKFNQALLVRYTKVPQVST 444
QY 421 PTLVEVSRLNGKVGSKCKGHPKAKMPCAEEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRLNGKVGSKCKGHPKAKMPCAEEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 504
QY 481 LVNRRPFCFSALEVDETVVPKFEFNAETFTTHADICTLSEKEROIKKQTALVELVKHKPKAT 540
DB 505 LVNRRPFCFSALEVDETVVPKFEFNAETFTTHADICTLSEKEROIKKQTALVELVKHKPKAT 564
QY 541 KEQLKAWMDFAAFVEKCKCKADDDKTCFAEEGKKLVAAASQAALGL 585
```

265 VHTCCCHGDLLECADRADLAKYICENQDSISSKJKECCCKPILLEKSHGICAEVENDMPA 324
301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEVARRHPDYSVWLLRLAKTYETTLK 360
325 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEVARRHPDYSVWLLRLAKTYETTLK 384
361 CAAADPHECYAKVDFEKPFLVEEPQNLKQNCBFLQGEYKFNQALLVRYTKVPQVST 420
385 CAAADPHECYAKVDFEKPFLVEEPQNLKQNCBFLQGEYKFNQALLVRYTKVPQVST 444
421 PTLVESENLGKSGCKCKHPEAKMPCAEADYLSVNLQCVLHEKTPVSDRVTKCCTES 480
445 PTLVESENLGKSGCKCKHPEAKMPCAEADYLSVNLQCVLHEKTPVSDRVTKCCTES 504
481 LVNRPCFSALEVEDTVVPKEFNAETTFHADICTLSEKERQIKKQATLVELVKKPKAT 540
505 LVNRPCFSALEVEDTVVPKEFNAETTFHADICTLSEKERQIKKQATLVELVKKPKAT 564
541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKLVAAASQAALGL 585
565 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKLVAAASQAALGL 609

RESULT 31
US-09-984-186-2
; Sequence 2, Application US/09984186
; Patent No. US20020151011A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittou, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,186
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-984-186-2
Query Match 100.0%; Score 585; DB 9; Length 610;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRFKDLGBENFKALVLIIFAQYLOQCPEFHVKLVNEVTEPAKTCVADESAE 60
DB 25 DAHKSEVAHRFKDLGBENFKALVLIIFAQYLOQCPEFHVKLVNEVTEPAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNECFLOHKDDNPNLPRLVREPV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNECFLOHKDDNPNLPRLVREPV 144
QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFAPBELLFPAKRYKAAFTCCQAADKAACLLP 180
DB 145 DVMCTAFHDNEETFLKKYLYEIAARRHPYFAPBELLFPAKRYKAAFTCCQAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
DB 205 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 264
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKJKECCCKPILLEKSHGICAEVENDMPA 300
DB 265 VHTCCCHGDLLECADRADLAKYICENQDSISSKJKECCCKPILLEKSHGICAEVENDMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEVARRHPDYSVWLLRLAKTYETTLK 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEVARRHPDYSVWLLRLAKTYETTLK 384
QY 361 CAAADPHECYAKVDFEKPFLVEEPQNLKQNCBFLQGEYKFNQALLVRYTKVPQVST 420
DB 385 CAAADPHECYAKVDFEKPFLVEEPQNLKQNCBFLQGEYKFNQALLVRYTKVPQVST 444
QY 421 PTLVESENLGKSGCKCKHPEAKMPCAEADYLSVNLQCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVESENLGKSGCKCKHPEAKMPCAEADYLSVNLQCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRPCFSALEVEDTVVPKEFNAETTFHADICTLSEKERQIKKQATLVELVKKPKAT 540
DB 505 LVNRPCFSALEVEDTVVPKEFNAETTFHADICTLSEKERQIKKQATLVELVKKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKLVAAASQAALGL 585
DB 565 KEQLKAVMDDFAAFVEKCKCKADDKETCFABEGKLVAAASQAALGL 609

RESULT 32
US-10-237-667-2
; Sequence 2, Application US/10237667
; Publication No. US20030022308A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittou, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 145 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSORFFPKAEFAEVSCLVTDLT 240
DB 205 KLDELDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSORFFPKAEFAEVSCLVTDLT 264
QY 241 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 300
DB 145 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 324
QY 181 KLDELDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSORFFPKAEFAEVSCLVTDLT 240
DB 205 KLDELDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSORFFPKAEFAEVSCLVTDLT 264
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
DB 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 324
QY 301 DPLSLAADPVESKDVCKNVAEAKDVLGMLFYEYARRHPDYSVLLRLAKTYETTLK 360
DB 325 DPLSLAADPVESKDVCKNVAEAKDVLGMLFYEYARRHPDYSVLLRLAKTYETTLK 384
QY 361 CAADPHCYAKVDFEFKPLVEEPQNLKONCELFEOQGEYKFNQALLVRYTKKVPQVST 420
DB 385 CAADPHCYAKVDFEFKPLVEEPQNLKONCELFEOQGEYKFNQALLVRYTKKVPQVST 444
QY 421 PTLVEVSRLNGKVGSKCKGHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRLNGKVGSKCKGHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDYTVPKFNAETFTPHADICTLSEKEROIKKQATLALVELVXHKPKAT 540
DB 505 LVNRRPCFSALEVDYTVPKFNAETFTPHADICTLSEKEROIKKQATLALVELVXHKPKAT 564
QY 541 KEQLKAVMDDDFAAFVKECKCKADDDKTCFAEKGKLVAAASQAALGL 585
DB 565 KEQLKAVMDDDFAAFVKECKCKADDDKTCFAEKGKLVAAASQAALGL 609

RESULT 29

US-10-153-604A-7
; Sequence 7, Application US/10153604A
; Publication No. US20030143191A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: P5556
; CURRENT APPLICATION NUMBER: US/10/153,604A
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-604A-7

Query Match 100.0%; Score 585; DB 14; Length 609;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144

QY 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 145 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSORFFPKAEFAEVSCLVTDLT 240
DB 205 KLDELDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSORFFPKAEFAEVSCLVTDLT 264
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
DB 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 324
QY 301 DPLSLAADPVESKDVCKNVAEAKDVLGMLFYEYARRHPDYSVLLRLAKTYETTLK 360
DB 325 DPLSLAADPVESKDVCKNVAEAKDVLGMLFYEYARRHPDYSVLLRLAKTYETTLK 384
QY 361 CAADPHCYAKVDFEFKPLVEEPQNLKONCELFEOQGEYKFNQALLVRYTKKVPQVST 420
DB 385 CAADPHCYAKVDFEFKPLVEEPQNLKONCELFEOQGEYKFNQALLVRYTKKVPQVST 444
QY 421 PTLVEVSRLNGKVGSKCKGHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRLNGKVGSKCKGHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDYTVPKFNAETFTPHADICTLSEKEROIKKQATLALVELVXHKPKAT 540
DB 505 LVNRRPCFSALEVDYTVPKFNAETFTPHADICTLSEKEROIKKQATLALVELVXHKPKAT 564
QY 541 KEQLKAVMDDDFAAFVKECKCKADDDKTCFAEKGKLVAAASQAALGL 585
DB 565 KEQLKAVMDDDFAAFVKECKCKADDDKTCFAEKGKLVAAASQAALGL 609

RESULT 30

US-10-365-623-23
; Sequence 23, Application US/10365623
; Publication No. US20030166512A1
; GENERAL INFORMATION:
; APPLICANT: Xie, Dong
; TITLE OF INVENTION: Protein Carrier System for Therapeutic Oligonucleotides
; FILE REFERENCE: 63024.000001
; CURRENT APPLICATION NUMBER: US/10/365,623
; CURRENT FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-365-623-23

Query Match 100.0%; Score 585; DB 14; Length 609;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 145 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSORFFPKAEFAEVSCLVTDLT 240
DB 205 KLDELDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSORFFPKAEFAEVSCLVTDLT 264
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300

```
/ TYPE: amino acid
/ STRANDEDNESS: <Unknown>
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-984-010-7

Query Match      100.0%; Score 585; DB 10; Length 604;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRRFKDLGEENFKALVLIAPQYLOQCPPEHVKLVNEVTEFAKTCVADESAB 60
DB 20 DAHKEVAHRRFKDLGEENFKALVLIAPQYLOQCPPEHVKLVNEVTEFAKTCVADESAB 79
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 80 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 139
QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKYKAAFTCCQAAADKAACLIP 180
DB 140 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKYKAAFTCCQAAADKAACLIP 199
QY 181 KLDELDRDEGKASSAKQRLKCSLOKFGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTK 240
DB 200 KLDELDRDEGKASSAKQRLKCSLOKFGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTK 259
QY 241 VHTCCCHGDLLECADRRADLAKYICENQDSISISKLKECCCKPILLESKSHCIAEVENDEMPA 300
DB 260 VHTCCCHGDLLECADRRADLAKYICENQDSISISKLKECCCKPILLESKSHCIAEVENDEMPA 319
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVVLRLRLAKTYETTLTK 360
DB 320 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVVLRLRLAKTYETTLTK 379
QY 361 CAADAPHECYAKVDFDFKPLVEEPQNLIKQNCSELFQELGEYKFNALLVRYTKVPQVST 420
DB 380 CAADAPHECYAKVDFDFKPLVEEPQNLIKQNCSELFQELGEYKFNALLVRYTKVPQVST 439
QY 421 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 440 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 499
QY 481 LVNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 540
DB 500 LVNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 559
QY 541 KEQLKAVMDDFAAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
DB 560 KEQLKAVMDDFAAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 604

RESULT 27
US-09-919-039-370
; Sequence 370, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; PRIOR FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 370
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

Query Match      100.0%; Score 585; DB 13; Length 609;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. US20030108871A1 088957CD1
US-09-919-039-370

Query Match      100.0%; Score 585; DB 10; Length 609;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRRFKDLGEENFKALVLIAPQYLOQCPPEHVKLVNEVTEFAKTCVADESAB 60
DB 25 DAHKEVAHRRFKDLGEENFKALVLIAPQYLOQCPPEHVKLVNEVTEFAKTCVADESAB 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKYKAAFTCCQAAADKAACLIP 180
DB 145 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKYKAAFTCCQAAADKAACLIP 204
QY 181 KLDELDRDEGKASSAKQRLKCSLOKFGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTK 240
DB 205 KLDELDRDEGKASSAKQRLKCSLOKFGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTK 264
QY 241 VHTCCCHGDLLECADRRADLAKYICENQDSISISKLKECCCKPILLESKSHCIAEVENDEMPA 300
DB 265 VHTCCCHGDLLECADRRADLAKYICENQDSISISKLKECCCKPILLESKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVVLRLRLAKTYETTLTK 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVVLRLRLAKTYETTLTK 384
QY 361 CAADAPHECYAKVDFDFKPLVEEPQNLIKQNCSELFQELGEYKFNALLVRYTKVPQVST 420
DB 385 CAADAPHECYAKVDFDFKPLVEEPQNLIKQNCSELFQELGEYKFNALLVRYTKVPQVST 444
QY 421 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 540
DB 505 LVNRRPCFSALEVDVETVYPKEFNAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 564
QY 541 KEQLKAVMDDFAAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
DB 565 KEQLKAVMDDFAAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 609

RESULT 28
US-10-153-064-7
; Sequence 7, Application US/10153064
; Publication No. US20020142814A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:

Query Match      100.0%; Score 585; DB 13; Length 609;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OTHER INFORMATION: human derived fusion protein

US-10-233-675A-11

Query Match 100.0%; Score 585; DB 15; Length 585;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDGLGEENFKALVLIARAFYLOQCFFEDHVKLVNEVTEFAKTCVADESAE 60
 DB 1 DAHKEVAHRFKDGLGEENFKALVLIARAFYLOQCFFEDHVKLVNEVTEFAKTCVADESAE 60
 QY 61 NCDKSLHTLFGDKLCTVATILRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATILRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 QY 121 DVMTAFHDNEETFLKKLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
 DB 121 DVMTAFHDNEETFLKKLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
 QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSQRPFAKFAEVSKLVTDLTK 240
 DB 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSQRPFAKFAEVSKLVTDLTK 240
 QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCXPLEKSHCIAEVENDMPA 300
 DB 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCXPLEKSHCIAEVENDMPA 300
 QY 301 DLPSLAADPVESKDVCKNYAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
 DB 301 DLPSLAADPVESKDVCKNYAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
 QY 361 CAADPHCYAKVDFBPKLVEBPQNLKQNCFLPEQLGEYKFNALLVRYTKVPQVST 420
 DB 361 CAADPHCYAKVDFBPKLVEBPQNLKQNCFLPEQLGEYKFNALLVRYTKVPQVST 420
 QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 QY 481 LVNRRPCFSALEVDYTPVPEFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540
 DB 481 LVNRRPCFSALEVDYTPVPEFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540
 QY 541 KEQLKAVMDDFAAVFEKCKCKADKTCFAEKGKLVAAASQAALGL 585
 DB 541 KEQLKAVMDDFAAVFEKCKCKADKTCFAEKGKLVAAASQAALGL 585
 RESULT 25
 US-10-462-262-26
 ; Sequence 26, Application US/10462262
 ; Publication No. US20040009534A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sato, Aaron K.
 ; TITLE OF INVENTION: PROTEIN ANALYSIS
 ; FILE REFERENCE: 10280-052001
 ; CURRENT APPLICATION NUMBER: US/10/462,262
 ; PRIOR FILING DATE: 2003-06-16
 ; PRIOR APPLICATION NUMBER: US 60/388,642
 ; NUMBER OF SEQ ID NOS: 430
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 26
 ; LENGTH: 585
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-462-262-26
 Query Match 100.0%; Score 585; DB 15; Length 585;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDGLGEENFKALVLIARAFYLOQCFFEDHVKLVNEVTEFAKTCVADESAE 60
 DB 1 DAHKEVAHRFKDGLGEENFKALVLIARAFYLOQCFFEDHVKLVNEVTEFAKTCVADESAE 60
 QY 61 NCDKSLHTLFGDKLCTVATILRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATILRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 QY 121 DVMTAFHDNEETFLKKLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
 DB 121 DVMTAFHDNEETFLKKLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
 QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSQRPFAKFAEVSKLVTDLTK 240
 DB 181 KLDELDEGKASSAKORLKCASLOKFGERAFKAWAVARLSQRPFAKFAEVSKLVTDLTK 240
 QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCXPLEKSHCIAEVENDMPA 300
 DB 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCXPLEKSHCIAEVENDMPA 300
 QY 301 DLPSLAADPVESKDVCKNYAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
 DB 301 DLPSLAADPVESKDVCKNYAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
 QY 361 CAADPHCYAKVDFBPKLVEBPQNLKQNCFLPEQLGEYKFNALLVRYTKVPQVST 420
 DB 361 CAADPHCYAKVDFBPKLVEBPQNLKQNCFLPEQLGEYKFNALLVRYTKVPQVST 420
 QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 QY 481 LVNRRPCFSALEVDYTPVPEFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540
 DB 481 LVNRRPCFSALEVDYTPVPEFNAETFTFHADICTLSEKERQIKKQTALVELVKKPKAT 540
 QY 541 KEQLKAVMDDFAAVFEKCKCKADKTCFAEKGKLVAAASQAALGL 585
 DB 541 KEQLKAVMDDFAAVFEKCKCKADKTCFAEKGKLVAAASQAALGL 585

RESULT 26

US-09-984-010-7
 ; Sequence 7, Application US/09984010
 ; Publication No. US20030104578A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ballance, David James
 ; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
 ; NUMBER OF SEQUENCES: 26
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
 ; STREET: 1300 I Street, NW
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20005-3315
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/984,010
 ; FILING DATE: 21-May-2002
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 09/091,873
 ; FILING DATE: 25-JUN-1998
 ; APPLICATION NUMBER: PCT/GB96/03164
 ; FILING DATE: 19-DEC-1996
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 604 amino acids

QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELDEGKASSAKORLKASQKFGGERAFKAWAVARLSORFPKABFAVSKLVTDLT 240
DB 181 KLDELDEGKASSAKORLKASQKFGGERAFKAWAVARLSORFPKABFAVSKLVTDLT 240
QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 300
DB 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHHPDYVSVLLRLAKTVETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHHPDYVSVLLRLAKTVETTLK 360
QY 361 CAAADPHECYAKVDFEFPKPLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKVPQVST 420
DB 361 CAAADPHECYAKVDFEFPKPLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGVSKCKCKHPEAKMPCAEEDYLSVNLQCLVLEHKTTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGVSKCKCKHPEAKMPCAEEDYLSVNLQCLVLEHKTTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDVETVYKPEFNAETFFHADICTLSEKERQIKKQATLVELVKKHKPAT 540
DB 481 LVNRRPCFSALEVDVETVYKPEFNAETFFHADICTLSEKERQIKKQATLVELVKKHKPAT 540
QY 541 KEQLKAVMDDFAAAFVEKCKCKADKCTCFABEGKGLVAASQAALGL 585
DB 541 KEQLKAVMDDFAAAFVEKCKCKADKCTCFABEGKGLVAASQAALGL 585

RESULT 21
US-10-413-832-2
; Sequence 2, Application US/10413832
; Publication No. US20030215359A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISC007
; CURRENT APPLICATION NUMBER: US/10/413,832
; PRIOR FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: US/09/806,247
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1)..(585)
; OTHER INFORMATION: ACETYLTATION
US-10-413-832-2

Query Match 100.0%; Score 585; DB 15; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGBENFKALVLIAPAYLOQCPEDHVKLVNEVTEFAKTCVADESAB 60
DB 1 DAHSEVAHRFKDLGBENFKALVLIAPAYLOQCPEDHVKLVNEVTEFAKTCVADESAB 60
QY 61 NCDLSLHTLFGDKLCTVATIRETYGEMADCCAKQEPERNECFLOHDDNPNLPRLVRPEV 120
DB 61 NCDLSLHTLFGDKLCTVATIRETYGEMADCCAKQEPERNECFLOHDDNPNLPRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELDEGKASSAKORLKASQKFGGERAFKAWAVARLSORFPKABFAVSKLVTDLT 240
DB 181 KLDELDEGKASSAKORLKASQKFGGERAFKAWAVARLSORFPKABFAVSKLVTDLT 240
QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 300
DB 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHHPDYVSVLLRLAKTVETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHHPDYVSVLLRLAKTVETTLK 360
QY 361 CAAADPHECYAKVDFEFPKPLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKVPQVST 420
DB 361 CAAADPHECYAKVDFEFPKPLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGVSKCKCKHPEAKMPCAEEDYLSVNLQCLVLEHKTTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGVSKCKCKHPEAKMPCAEEDYLSVNLQCLVLEHKTTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDVETVYKPEFNAETFFHADICTLSEKERQIKKQATLVELVKKHKPAT 540
DB 481 LVNRRPCFSALEVDVETVYKPEFNAETFFHADICTLSEKERQIKKQATLVELVKKHKPAT 540
QY 541 KEQLKAVMDDFAAAFVEKCKCKADKCTCFABEGKGLVAASQAALGL 585
DB 541 KEQLKAVMDDFAAAFVEKCKCKADKCTCFABEGKGLVAASQAALGL 585

RESULT 22
US-10-414-386-1
; Sequence 1, Application US/10414386
; Publication No. US20030215952A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISC007
; CURRENT APPLICATION NUMBER: US/10/414,386
; PRIOR FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: US/09/806,247
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-414-386-1

Query Match 100.0%; Score 585; DB 15; Length 585;

Db 241 VHTCCGDLLECADRADLAKYICENQDSISSKLECECKPILLEKSHGICIAEVENDMPA 300
Qy 301 DLPSLAADPVESKOVCKNYAKADVFLGMFLYEVARRHPDYSVWLLRLAKTYETTTLEKC 360
Db 301 DLPSLAADPVESKOVCKNYAKADVFLGMFLYEVARRHPDYSVWLLRLAKTYETTTLEKC 360
Qy 361 CAADPHECYAKVDFEFKPLVEEPQNLKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
Db 361 CAADPHECYAKVDFEFKPLVEEPQNLKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
Qy 421 PTLVEVSRLNKGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRVTKCTES 480
Db 421 PTLVEVSRLNKGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRVTKCTES 480
Qy 481 LVNRRPCFSALEVDVETYPKFNATFTFHADICTLSEKEROIKKOTALVELVKHKPKAT 540
Db 481 LVNRRPCFSALEVDVETYPKFNATFTFHADICTLSEKEROIKKOTALVELVKHKPKAT 540
Qy 541 KEQLKAVMDDFAAFVEKCKADDKTCFAEKGKLVAAASQAALGL 585
Db 541 KEQLKAVMDDFAAFVEKCKADDKTCFAEKGKLVAAASQAALGL 585
RESULT 19
US-10-413-831-2
; Sequence 2, Application US/10413831
; Publication No. US20030194813A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISCO07
; CURRENT APPLICATION NUMBER: US/10/413,831
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: US/09/806,247
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1)..(585)
; OTHER INFORMATION: ACETYLTATION
US-10-413-831-2
Query Match 100.0%; Score 585; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DAHKSEVAHRFKDLGGENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESA 60
Db 1 DAHKSEVAHRFKDLGGENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESA 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLOHKDNDPNLPRVPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLOHKDNDPNLPRVPEV 120
Qy 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFVAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180

Db 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFVAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
Qy 181 KLDELDEGRASSAKQRLKCSAQKQGERAFKAWAVARLSQRPFAEFAVSKLVTDLT 240
Db 181 KLDELDEGRASSAKQRLKCSAQKQGERAFKAWAVARLSQRPFAEFAVSKLVTDLT 240
Qy 241 VHTCCGDLLECADRADLAKYICENQDSISSKLECECKPILLEKSHGICIAEVENDMPA 300
Db 241 VHTCCGDLLECADRADLAKYICENQDSISSKLECECKPILLEKSHGICIAEVENDMPA 300
Qy 301 DLPSLAADPVESKOVCKNYAKADVFLGMFLYEVARRHPDYSVWLLRLAKTYETTTLEKC 360
Db 301 DLPSLAADPVESKOVCKNYAKADVFLGMFLYEVARRHPDYSVWLLRLAKTYETTTLEKC 360
Qy 361 CAADPHECYAKVDFEFKPLVEEPQNLKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
Db 361 CAADPHECYAKVDFEFKPLVEEPQNLKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
Qy 421 PTLVEVSRLNKGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRVTKCTES 480
Db 421 PTLVEVSRLNKGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRVTKCTES 480
Qy 481 LVNRRPCFSALEVDVETYPKFNATFTFHADICTLSEKEROIKKOTALVELVKHKPKAT 540
Db 481 LVNRRPCFSALEVDVETYPKFNATFTFHADICTLSEKEROIKKOTALVELVKHKPKAT 540
Qy 541 KEQLKAVMDDFAAFVEKCKADDKTCFAEKGKLVAAASQAALGL 585
Db 541 KEQLKAVMDDFAAFVEKCKADDKTCFAEKGKLVAAASQAALGL 585
RESULT 20
US-10-413-832-1
; Sequence 1, Application US/10413832
; Publication No. US20030215359A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISCO07
; CURRENT APPLICATION NUMBER: US/10/413,832
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: US/09/806,247
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-413-832-1
Query Match 100.0%; Score 585; DB 15; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DAHKSEVAHRFKDLGGENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESA 60
Db 1 DAHKSEVAHRFKDLGGENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESA 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLOHKDNDPNLPRVPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLOHKDNDPNLPRVPEV 120

Db 421 PTLVEVSRNLGVSKCKHPEAKRMPCAEYLSVVLNQLCVLHEKTPVSDRVTKCTTES 480
QY 481 LVNRAPCFSALEVDYVPKFNATFTFHADICTLSEKEROIKKQTALVELVKHKPKAT 540
Db 481 LVNRAPCFSALEVDYVPKFNATFTFHADICTLSEKEROIKKQTALVELVKHKPKAT 540
QY 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEKGKLVAAQAALGL 585
Db 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEKGKLVAAQAALGL 585

RESULT 17
US-10-414-469-2
; Sequence 2, Application US/10414469
; Publication No. US20030190691A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISCO07
; CURRENT APPLICATION NUMBER: US/10/414,469
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: 09/806,247
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: PCT/US99/22905
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (1)..(585)
; OTHER INFORMATION: ACETYLTATION
US-10-414-469-2

Query Match 100.0%; Score 585; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHREFKDLGEENFKALVLIAPAYLQOCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHREFKDLGEENFKALVLIAPAYLQOCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
QY 121 DVMTCTAFHNEETFLKKYLYEIAARRHPFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
Db 121 DVMTCTAFHNEETFLKKYLYEIAARRHPFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFPKAEFAEVSCLVTDLTJK 240
Db 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFPKAEFAEVSCLVTDLTJK 240
QY 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300

QY 301 DLPSLAADFVESKDVCKNVAEAKOVFLGMFLYEHARRHPDYSVLLLLAKTYETTLK 360
Db 301 DLPSLAADFVESKDVCKNVAEAKOVFLGMFLYEHARRHPDYSVLLLLAKTYETTLK 360
QY 361 CAAADPHECVAKVDFEFKPLVEEPQNLIKONCELPEQLGEYKFNALLVRYTKVPQVST 420
Db 361 CAAADPHECVAKVDFEFKPLVEEPQNLIKONCELPEQLGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGVSKCKHPEAKRMPCAEYLSVVLNQLCVLHEKTPVSDRVTKCTTES 480
Db 421 PTLVEVSRNLGVSKCKHPEAKRMPCAEYLSVVLNQLCVLHEKTPVSDRVTKCTTES 480
QY 481 LVNRAPCFSALEVDYVPKFNATFTFHADICTLSEKEROIKKQTALVELVKHKPKAT 540
Db 481 LVNRAPCFSALEVDYVPKFNATFTFHADICTLSEKEROIKKQTALVELVKHKPKAT 540
QY 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEKGKLVAAQAALGL 585
Db 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEKGKLVAAQAALGL 585

RESULT 18
US-10-413-831-1
; Sequence 1, Application US/10413831
; Publication No. US20030194813A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISCO07
; CURRENT APPLICATION NUMBER: US/10/413,831
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: US/09/806,247
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-413-831-1

Query Match 100.0%; Score 585; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHREFKDLGEENFKALVLIAPAYLQOCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHREFKDLGEENFKALVLIAPAYLQOCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
QY 121 DVMTCTAFHNEETFLKKYLYEIAARRHPFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
Db 121 DVMTCTAFHNEETFLKKYLYEIAARRHPFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFPKAEFAEVSCLVTDLTJK 240
Db 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFPKAEFAEVSCLVTDLTJK 240
QY 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300

```

; Sequence 2, Application US/10319263
; Publication No. US20030180820A1
; GENERAL INFORMATION:
; APPLICANT: Lau Ph.D., David
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: IS0007
; CURRENT APPLICATION NUMBER: US/10/319,263
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (1)..(585)
; OTHER INFORMATION: ACETYLATION
US-10-319-263-2

Query Match 100.0%; Score 585; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVLIATFAQYLOQCPFDHVKLVNEVTEFAKTCVADSEAE 60
DB 1 DAHKEVAHRFKDLGEENFKALVLIATFAQYLOQCPFDHVKLVNEVTEFAKTCVADSEAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
QY 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
DB 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
QY 181 KLDELDEGKASSAKORLKASLOKGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTK 240
DB 181 KLDELDEGKASSAKORLKASLOKGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTK 240
QY 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLKCECKEPLEKSHCIAEVENDEMPA 300
DB 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLKCECKEPLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVSKDVCNKAEDKDFLGMLFYEARHPDYVSVLLLR LAKTYETTTLEK 360
DB 301 DLPSLAADFVSKDVCNKAEDKDFLGMLFYEARHPDYVSVLLLR LAKTYETTTLEK 360
QY 361 CAAADPHECVAKVDFEFLVEEPQNLKQNCLEFEQLGEYKFQNALIVRYTKKVPQVST 420
DB 361 CAAADPHECVAKVDFEFLVEEPQNLKQNCLEFEQLGEYKFQNALIVRYTKKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKCHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKCHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETVVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540
DB 481 LVNRRPCFSALEVDETVVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540

; Sequence 1, Application US/10414469
; Publication No. US20030190691A1
; GENERAL INFORMATION:
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: IS0007
; CURRENT APPLICATION NUMBER: US/10/414,469
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: 09/806,247
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: PCT/US99/22905
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-414-469-1

Query Match 100.0%; Score 585; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGEENFKALVLIATFAQYLOQCPFDHVKLVNEVTEFAKTCVADSEAE 60
DB 1 DAHKEVAHRFKDLGEENFKALVLIATFAQYLOQCPFDHVKLVNEVTEFAKTCVADSEAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
QY 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
DB 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
QY 181 KLDELDEGKASSAKORLKASLOKGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTK 240
DB 181 KLDELDEGKASSAKORLKASLOKGERAFKAWAVARLSORFPKAEFAEVS KLVTDLTK 240
QY 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLKCECKEPLEKSHCIAEVENDEMPA 300
DB 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLKCECKEPLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVSKDVCNKAEDKDFLGMLFYEARHPDYVSVLLLR LAKTYETTTLEK 360
DB 301 DLPSLAADFVSKDVCNKAEDKDFLGMLFYEARHPDYVSVLLLR LAKTYETTTLEK 360
QY 361 CAAADPHECVAKVDFEFLVEEPQNLKQNCLEFEQLGEYKFQNALIVRYTKKVPQVST 420
DB 361 CAAADPHECVAKVDFEFLVEEPQNLKQNCLEFEQLGEYKFQNALIVRYTKKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKCHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKCHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETVVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540
DB 481 LVNRRPCFSALEVDETVVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540
```

QY 481 LVNRRPCFSALEVDYVPEKFNAAETTFTHADICTLSEKERQIKKOTALVELVGHKPKAT 540
Db 481 LVNRRPCFSALEVDYVPEKFNAAETTFTHADICTLSEKERQIKKOTALVELVGHKPKAT 540
QY 541 KEOLKAVMDDFAAFVEKCKCKADDDKTCFAEBGKCLVAASQAALGL 585
Db 541 KEOLKAVMDDFAAFVEKCKCKADDDKTCFAEBGKCLVAASQAALGL 585

RESULT 13
US-10-153-604A-5
; Sequence 5, Application US/10153604A
; Publication No. US20030143191A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; CURRENT APPLICATION NUMBER: US/10/153,604A
; FILE REFERENCE: PF556
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-604A-5

Query Match 100.0%; Score 585; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
Db 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMTAFHNDNEFTLKKLYEIAIRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 121 DVMTAFHNDNEFTLKKLYEIAIRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPFAEVSCLVTDLT 240
Db 181 KLDELDEGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPFAEVSCLVTDLT 240
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDempa 300
Db 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDempa 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
QY 361 CAAADPHCYAKVDFEFPKPLVEEPQNLIKONCELFQGLGEYKFNALLVRYTKVPQVST 420
Db 361 CAAADPHCYAKVDFEFPKPLVEEPQNLIKONCELFQGLGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSNLKGKSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSNLKGKSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDYVPEKFNAAETTFTHADICTLSEKERQIKKOTALVELVGHKPKAT 540
Db 481 LVNRRPCFSALEVDYVPEKFNAAETTFTHADICTLSEKERQIKKOTALVELVGHKPKAT 540
QY 541 KEOLKAVMDDFAAFVEKCKCKADDDKTCFAEBGKCLVAASQAALGL 585
Db 541 KEOLKAVMDDFAAFVEKCKCKADDDKTCFAEBGKCLVAASQAALGL 585

RESULT 14
US-10-319-263-1
; Sequence 1, Application US/10319263
; Publication No. US20030180820A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISC007
; CURRENT APPLICATION NUMBER: US/10/319,263
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-319-263-1

Query Match 100.0%; Score 585; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
Db 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMTAFHNDNEFTLKKLYEIAIRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 121 DVMTAFHNDNEFTLKKLYEIAIRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPFAEVSCLVTDLT 240
Db 181 KLDELDEGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPFAEVSCLVTDLT 240
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDempa 300
Db 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDempa 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
QY 361 CAAADPHCYAKVDFEFPKPLVEEPQNLIKONCELFQGLGEYKFNALLVRYTKVPQVST 420
Db 361 CAAADPHCYAKVDFEFPKPLVEEPQNLIKONCELFQGLGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSNLKGKSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSNLKGKSKCKCHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDYVPEKFNAAETTFTHADICTLSEKERQIKKOTALVELVGHKPKAT 540
Db 481 LVNRRPCFSALEVDYVPEKFNAAETTFTHADICTLSEKERQIKKOTALVELVGHKPKAT 540
QY 541 KEOLKAVMDDFAAFVEKCKCKADDDKTCFAEBGKCLVAASQAALGL 585
Db 541 KEOLKAVMDDFAAFVEKCKCKADDDKTCFAEBGKCLVAASQAALGL 585

Db 301 DLPSLAADFVESKDVCKNKAARAKVFLGMFLYEYARRHPDYSVLLLRLLAKTYETTTLEK 360
QY 361 CAADPHECYAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420
Db 361 CAADPHECYAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
Db 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
QY 481 LVNRRPCFSALEVDYVYPKEFNAETFTFHADICTLSEKEROIKKQTALVELVGHKPKAT 540
Db 481 LVNRRPCFSALEVDYVYPKEFNAETFTFHADICTLSEKEROIKKQTALVELVGHKPKAT 540
QY 541 KEOLKAVMDDFAAFVCKCKADDDKTCFAEBEGKLVAAASQAALGL 585
Db 541 KEOLKAVMDDFAAFVCKCKADDDKTCFAEBEGKLVAAASQAALGL 585

RESULT 11
US-10-425-000-31
; Sequence 31, Application US/10425000
; Publication No. US2004005277A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
; TITLE OF INVENTION: Angiogenesis
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-31

Query Match 100.0%; Score 585; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPEDHVKLVNEVTEPAKTCVADESAB 60
Db 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPEDHVKLVNEVTEPAKTCVADESAB 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPFLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPFLRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFAPBELLFFAKRYKAATFECQQAADKAACLLP 180
Db 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFAPBELLFFAKRYKAATFECQQAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
Db 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
QY 241 VHTTECHGDLLECCADRDADIAKYICENQDSISKLKECKEPLLEKSHCIAEVENDEMPA 300
Db 241 VHTTECHGDLLECCADRDADIAKYICENQDSISKLKECKEPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNKAARAKVFLGMFLYEYARRHPDYSVLLLRLLAKTYETTTLEK 360
Db 301 DLPSLAADFVESKDVCKNKAARAKVFLGMFLYEYARRHPDYSVLLLRLLAKTYETTTLEK 360

QY 361 CAADPHECYAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420
Db 361 CAADPHECYAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
Db 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
QY 481 LVNRRPCFSALEVDYVYPKEFNAETFTFHADICTLSEKEROIKKQTALVELVGHKPKAT 540
Db 481 LVNRRPCFSALEVDYVYPKEFNAETFTFHADICTLSEKEROIKKQTALVELVGHKPKAT 540
QY 541 KEOLKAVMDDFAAFVCKCKADDDKTCFAEBEGKLVAAASQAALGL 585
Db 541 KEOLKAVMDDFAAFVCKCKADDDKTCFAEBEGKLVAAASQAALGL 585

RESULT 12
US-10-153-064-5
; Sequence 5, Application US/10153064
; Publication No. US2002014281A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5

Query Match 100.0%; Score 585; DB 13; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPEDHVKLVNEVTEPAKTCVADESAB 60
Db 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLQCCPEDHVKLVNEVTEPAKTCVADESAB 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPFLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPFLRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFAPBELLFFAKRYKAATFECQQAADKAACLLP 180
Db 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFAPBELLFFAKRYKAATFECQQAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
Db 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
QY 241 VHTTECHGDLLECCADRDADIAKYICENQDSISKLKECKEPLLEKSHCIAEVENDEMPA 300
Db 241 VHTTECHGDLLECCADRDADIAKYICENQDSISKLKECKEPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNKAARAKVFLGMFLYEYARRHPDYSVLLLRLLAKTYETTTLEK 360
Db 301 DLPSLAADFVESKDVCKNKAARAKVFLGMFLYEYARRHPDYSVLLLRLLAKTYETTTLEK 360
QY 361 CAADPHECYAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420
Db 361 CAADPHECYAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
Db 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480

Db 181 KDELDEGKASSAKQRLKCLASQKQGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 240
QY 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLKECKEPLLEKSHCIAEVENDEMPA 300
Db 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLKECKEPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLRLAKTYETTLTKC 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLRLAKTYETTLTKC 360
QY 361 CAADPHCEYAKVDFEFPKPLVEBPQNLIKQNCLEPEQIGYKFNQALLVRYTKKVPQVST 420
Db 361 CAADPHCEYAKVDFEFPKPLVEBPQNLIKQNCLEPEQIGYKFNQALLVRYTKKVPQVST 420
QY 421 PTLVSVSNLGVKSGCKCKHPEAKRMPCAEDYLSVNLQCLVHEKTPVSDRVTKCCTES 480
Db 421 PTLVSVSNLGVKSGCKCKHPEAKRMPCAEDYLSVNLQCLVHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDYTVPKFNAETFTFHADICTLSEKEROIKQTALVELVKKHKKPAT 540
Db 481 LVNRRPCFSALEVDYTVPKFNAETFTFHADICTLSEKEROIKQTALVELVKKHKKPAT 540
QY 541 KEQLKAVMDDFAAFEVKCKCKADDKTCFAEKGKLVAAASQAALGL 585
Db 541 KEQLKAVMDDFAAFEVKCKCKADDKTCFAEKGKLVAAASQAALGL 585

RESULT 9
US-09-833-245-18
; Sequence 18, Application US/09833245
; Publication No. US200400101341
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 18
; TYPE: PRT
; LENGTH: 585
; ORGANISM: Homo Sapiens
US-09-833-245-18

Query Match 100.0%; Score 585; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
Db 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
QY 181 KDELDEGKASSAKQRLKCLASQKQGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 240
Db 181 KDELDEGKASSAKQRLKCLASQKQGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 240
QY 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLKECKEPLLEKSHCIAEVENDEMPA 300

Db 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLKECKEPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLRLAKTYETTLTKC 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLRLAKTYETTLTKC 360
QY 361 CAADPHCEYAKVDFEFPKPLVEBPQNLIKQNCLEPEQIGYKFNQALLVRYTKKVPQVST 420
Db 361 CAADPHCEYAKVDFEFPKPLVEBPQNLIKQNCLEPEQIGYKFNQALLVRYTKKVPQVST 420
QY 421 PTLVSVSNLGVKSGCKCKHPEAKRMPCAEDYLSVNLQCLVHEKTPVSDRVTKCCTES 480
Db 421 PTLVSVSNLGVKSGCKCKHPEAKRMPCAEDYLSVNLQCLVHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDYTVPKFNAETFTFHADICTLSEKEROIKQTALVELVKKHKKPAT 540
Db 481 LVNRRPCFSALEVDYTVPKFNAETFTFHADICTLSEKEROIKQTALVELVKKHKKPAT 540
QY 541 KEQLKAVMDDFAAFEVKCKCKADDKTCFAEKGKLVAAASQAALGL 585
Db 541 KEQLKAVMDDFAAFEVKCKCKADDKTCFAEKGKLVAAASQAALGL 585

Query Match 100.0%; Score 585; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLOQCPFDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
Db 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLLP 180
QY 181 KDELDEGKASSAKQRLKCLASQKQGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 240
Db 181 KDELDEGKASSAKQRLKCLASQKQGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 240
QY 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLKECKEPLLEKSHCIAEVENDEMPA 300
Db 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKLKECKEPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLRLAKTYETTLTKC 360

181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240
181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240
241 VHTCCGHDLLCADDADLAKYICENQDSISSKLKECCCKPILLESKHGICIAEVENDEMPA 300
241 VHTCCGHDLLCADDADLAKYICENQDSISSKLKECCCKPILLESKHGICIAEVENDEMPA 300
301 DLPSLAADPVESKDVCKNYAEAKDVFLGMFLYEAARRHPDYSVLLLRLLAKTYETTLK 360
301 DLPSLAADPVESKDVCKNYAEAKDVFLGMFLYEAARRHPDYSVLLLRLLAKTYETTLK 360
361 CAADPHECYAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKVPQVST 420
361 CAADPHECYAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKVPQVST 420
421 PTLVEVSRLNGKVGSKCCCKHPKAPCAEDYLSVNLQCVLHEKTPVSDRVTKCCTES 480
421 PTLVEVSRLNGKVGSKCCCKHPKAPCAEDYLSVNLQCVLHEKTPVSDRVTKCCTES 480
481 LVNRRPCFSALEVDYVVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKHKPKAT 540
481 LVNRRPCFSALEVDYVVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKHKPKAT 540
541 KEOLKAVMDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
541 KEOLKAVMDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585

RESULT 7
US-09-832-501-18
; Sequence 18, Application US/09832501
; Publication No. US20030199043A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J.
; APPLICANT: Sleep, Darrell
; APPLICANT: Turner, Andrew J.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF542
; CURRENT APPLICATION NUMBER: US/09/832,501
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-832-501-18

Query Match 100.0%; Score 585; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADES 60
1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADES 60
61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120
61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120
121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFAPPELLFFAKRYKAATFCCQAADKAACLLP 180
121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFAPPELLFFAKRYKAATFCCQAADKAACLLP 180

181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240
181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240
241 VHTCCGHDLLCADDADLAKYICENQDSISSKLKECCCKPILLESKHGICIAEVENDEMPA 300
241 VHTCCGHDLLCADDADLAKYICENQDSISSKLKECCCKPILLESKHGICIAEVENDEMPA 300
301 DLPSLAADPVESKDVCKNYAEAKDVFLGMFLYEAARRHPDYSVLLLRLLAKTYETTLK 360
301 DLPSLAADPVESKDVCKNYAEAKDVFLGMFLYEAARRHPDYSVLLLRLLAKTYETTLK 360
361 CAADPHECYAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKVPQVST 420
361 CAADPHECYAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKVPQVST 420
421 PTLVEVSRLNGKVGSKCCCKHPKAPCAEDYLSVNLQCVLHEKTPVSDRVTKCCTES 480
421 PTLVEVSRLNGKVGSKCCCKHPKAPCAEDYLSVNLQCVLHEKTPVSDRVTKCCTES 480
481 LVNRRPCFSALEVDYVVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKHKPKAT 540
481 LVNRRPCFSALEVDYVVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKHKPKAT 540
541 KEOLKAVMDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
541 KEOLKAVMDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585

RESULT 8
US-09-833-118-18
; Sequence 18, Application US/09833118
; Publication No. US20030219875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF544
; CURRENT APPLICATION NUMBER: US/09/833,118
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-118-18

Query Match 100.0%; Score 585; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADES 60
1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADES 60
61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120
61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120
121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFAPPELLFFAKRYKAATFCCQAADKAACLLP 180
121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFAPPELLFFAKRYKAATFCCQAADKAACLLP 180
181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
DB 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
QY 181 KLDELDEGKASSAKORLKASLOKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
DB 181 KLDELDEGKASSAKORLKASLOKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSSAADFVESKDVCKNYAEAKDVLGMFLYVYARRHPDYSVLLLRLLAKTYETTTLEK 360
DB 301 DLPSSAADFVESKDVCKNYAEAKDVLGMFLYVYARRHPDYSVLLLRLLAKTYETTTLEK 360
QY 361 CAADPHECYAKVDFEPKPLVEEPQNLKQNCFLFEQYGEYKFNALLVYTKVPOVST 420
DB 361 CAADPHECYAKVDFEPKPLVEEPQNLKQNCFLFEQYGEYKFNALLVYTKVPOVST 420
QY 421 PTLVEVSRNLGKVGSKCKCKHPEAKMPCAEADYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKCKHPEAKMPCAEADYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVXHKPKAT 540
DB 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVXHKPKAT 540
QY 541 KEOLKAVMDDFAAFVEKCKCKADKCTCFABEGKKLVAASQAALGL 585
DB 541 KEOLKAVMDDFAAFVEKCKCKADKCTCFABEGKKLVAASQAALGL 585

RESULT 5
US-09-833-117-18
; Sequence 18, Application US/09833117
; Publication No. US20030171267A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF543
; CURRENT APPLICATION NUMBER: US/09/833,117
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-117-18

Query Match 100.0%; Score 585; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
DB 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
QY 181 KLDELDEGKASSAKORLKASLOKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
DB 181 KLDELDEGKASSAKORLKASLOKFGGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSSAADFVESKDVCKNYAEAKDVLGMFLYVYARRHPDYSVLLLRLLAKTYETTTLEK 360
DB 301 DLPSSAADFVESKDVCKNYAEAKDVLGMFLYVYARRHPDYSVLLLRLLAKTYETTTLEK 360
QY 361 CAADPHECYAKVDFEPKPLVEEPQNLKQNCFLFEQYGEYKFNALLVYTKVPOVST 420
DB 361 CAADPHECYAKVDFEPKPLVEEPQNLKQNCFLFEQYGEYKFNALLVYTKVPOVST 420
QY 421 PTLVEVSRNLGKVGSKCKCKHPEAKMPCAEADYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKCKHPEAKMPCAEADYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVXHKPKAT 540
DB 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVXHKPKAT 540
QY 541 KEOLKAVMDDFAAFVEKCKCKADKCTCFABEGKKLVAASQAALGL 585
DB 541 KEOLKAVMDDFAAFVEKCKCKADKCTCFABEGKKLVAASQAALGL 585

RESULT 6
US-09-932-322-445
; Sequence 445, Application US/09932322
; Publication No. US20030194743A1
; GENERAL INFORMATION:
; APPLICANT: Dyax Corp.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Ladner, Robert Charles
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BLYS)
; FILE REFERENCE: DYX-018.1 PCT; DYX-018.1 US
; CURRENT APPLICATION NUMBER: US/09/932,322
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-932-322-445

Query Match 100.0%; Score 585; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
DB 121 DVMCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180

Db 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECCADRADLAKYICENODSISSKLKECCCKPILLEKSHGICIAEVNDMPA 300
Db 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPILLEKSHGICIAEVNDMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEYARRHPDYSVVLRLRLAKTYETTLK 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEYARRHPDYSVVLRLRLAKTYETTLK 360
QY 361 CAADAPHECYAKVDFEFKPLVEEPQNLIKONCELFEQGEYKFNALLVRYTKVPQVST 420
Db 361 CAADAPHECYAKVDFEFKPLVEEPQNLIKONCELFEQGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALVDETVYPKEFNAETFTFHADICTLSEKEROIKKQTALVELVGHKPKAT 540
Db 481 LVNRRPCFSALVDETVYPKEFNAETFTFHADICTLSEKEROIKKQTALVELVGHKPKAT 540
QY 541 KEOLKAVMDDFAAFVEKCKKADDDKTCFAEGKGLVAASQAALGL 585
Db 541 KEOLKAVMDDFAAFVEKCKKADDDKTCFAEGKGLVAASQAALGL 585

RESULT 3

US-09-984-010-26
; Sequence 26, Application US/09984010
; Publication No. US20030104578A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David James
; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE AND SERUM ALBUMIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, PARABOW, GARRETT & DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,010
; FILING DATE: 21-May-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/091,873
; FILING DATE: 25-JUN-1998
; APPLICATION NUMBER: PCT/GB96/03164
; FILING DATE: 19-DEC-1996
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-984-010-26

Query Match 100.0%; Score 585; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPQYLOQCFFEDHVKLVNEVTEFAKTCVADESAE 60

Db 1 DAHKSEVAHRFKDLGEENFKALVLIAPQYLOQCFFEDHVKLVNEVTEFAKTCVADESAE 60
QY 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNIPLRVREV 120
Db 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNIPLRVREV 120
QY 121 DVNCTAPHDNEETFLKKYLVEIARRHPYFAPBELLFFAKRYKAAFTTECCQAADKAACLLP 180
Db 121 DVNCTAPHDNEETFLKKYLVEIARRHPYFAPBELLFFAKRYKAAFTTECCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
Db 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECCADRADLAKYICENODSISSKLKECCCKPILLEKSHGICIAEVNDMPA 300
Db 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPILLEKSHGICIAEVNDMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEYARRHPDYSVVLRLRLAKTYETTLK 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEYARRHPDYSVVLRLRLAKTYETTLK 360
QY 361 CAADAPHECYAKVDFEFKPLVEEPQNLIKONCELFEQGEYKFNALLVRYTKVPQVST 420
Db 361 CAADAPHECYAKVDFEFKPLVEEPQNLIKONCELFEQGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRLGKVGSKCKKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALVDETVYPKEFNAETFTFHADICTLSEKEROIKKQTALVELVGHKPKAT 540
Db 481 LVNRRPCFSALVDETVYPKEFNAETFTFHADICTLSEKEROIKKQTALVELVGHKPKAT 540
QY 541 KEOLKAVMDDFAAFVEKCKKADDDKTCFAEGKGLVAASQAALGL 585
Db 541 KEOLKAVMDDFAAFVEKCKKADDDKTCFAEGKGLVAASQAALGL 585

RESULT 4

US-09-833-041-18
; Sequence 18, Application US/09833041
; Publication No. US20030125247A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF545
; CURRENT APPLICATION NUMBER: US/09/833,041
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-041-18

Query Match 100.0%; Score 585; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPQYLOQCFFEDHVKLVNEVTEFAKTCVADESAE 60
Db 1 DAHKSEVAHRFKDLGEENFKALVLIAPQYLOQCFFEDHVKLVNEVTEFAKTCVADESAE 60

89 510 87.2 660 14 US-10-153-604A-90 Sequence 90, Appl
90 510 87.2 660 14 US-10-153-604A-93 Sequence 93, Appl
91 510 87.2 676 13 US-10-153-064-95 Sequence 95, Appl
92 510 87.2 676 13 US-10-153-064-98 Sequence 98, Appl
93 510 87.2 676 13 US-10-153-064-104 Sequence 104, Appl
94 510 87.2 676 14 US-10-153-604A-95 Sequence 95, Appl
95 510 87.2 676 14 US-10-153-604A-98 Sequence 98, Appl
96 510 87.2 676 14 US-10-153-604A-104 Sequence 104, Appl
97 510 87.2 684 13 US-10-153-064-92 Sequence 92, Appl
98 510 87.2 684 14 US-10-153-604A-92 Sequence 92, Appl
99 510 87.2 1184 13 US-10-153-064-89 Sequence 89, Appl
100 510 87.2 1184 14 US-10-153-604A-89 Sequence 89, Appl

ALIGNMENTS

RESULT 1

US-09-929-552-2
; Sequence 2, Application US/09929552
; Patent No. US20030123080A1
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; Soto, Ana M.

; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells

; CORRESPONDENCE ADDRESS: 2

; ADDRESS: Medlen & Carroll, LLP

; STREET: 220 Montgomery Street, Suite 2200

; CITY: San Francisco

; STATE: California

; COUNTRY: United States of America

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION NUMBER: US/09/929,552

; FILING DATE: 14-Aug-2001

; CLASSIFICATION: <Unknown>

; APPLICATION NUMBER: 08/769,746

; FILING DATE: 19-DEC-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Carroll, Peter G.

; REGISTRATION NUMBER: 32,837

; REFERENCE/DOCKET NUMBER: MERI-02584

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 705-8410

; TELEFAX: (415) 397-8338

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 585 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-929-552-2

Query Match 100.0%; Score 585; DB 9; Length 585;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHREKDLGSENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60

DB 1 DAHSEVAHREKDLGSENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRREV 120

DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRREV 120

QY 121 DVMTAFHNEETFLKKYLIEIARRHPFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180

DB 121 DVMTAFHNEETFLKKYLIEIARRHPFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180

QY 181 KLDELDEGKASSAKQRLKCAASLQKFGERAFAKAWAVARLSQRPFAEFAVSKLVTDLTK 240

DB 181 KLDELDEGKASSAKQRLKCAASLQKFGERAFAKAWAVARLSQRPFAEFAVSKLVTDLTK 240

QY 241 VHTECCHGDLLECADRADLAKYICENQDISISKLKECCKEPPLLEKSHCIAEVENDEMPA 300

DB 241 VHTECCHGDLLECADRADLAKYICENQDISISKLKECCKEPPLLEKSHCIAEVENDEMPA 300

QY 301 DLPSLAADFVESKDVCKNYAAEKDVFGLMFLYIYARRHPDYSVLLLRLLAKTYETTLEKC 360

DB 301 DLPSLAADFVESKDVCKNYAAEKDVFGLMFLYIYARRHPDYSVLLLRLLAKTYETTLEKC 360

QY 361 CAADPHCEYAKVDFEKLPIVEBPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420

DB 361 CAADPHCEYAKVDFEKLPIVEBPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420

QY 421 PTLVEVSRLNGKVGSKCCKHPEAKRMPCAEDYLSVVNLQICVLHEKTPVSDRVTKCCTES 480

DB 421 PTLVEVSRLNGKVGSKCCKHPEAKRMPCAEDYLSVVNLQICVLHEKTPVSDRVTKCCTES 480

QY 481 LVNRRPCFSALEVDVETVPKEFNAETTFHADICTLSEKEROIKKOTALVELVKKHKPKAT 540

DB 481 LVNRRPCFSALEVDVETVPKEFNAETTFHADICTLSEKEROIKKOTALVELVKKHKPKAT 540

QY 541 KEQLKAVMDDFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 585

DB 541 KEQLKAVMDDFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 585

RESULT 2

US-09-932-613-445
; Sequence 445, Application US/09932613
; Publication No. US20030091565A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON
; FILE REFERENCE: DYX-025.1 PCT; DYX-025.1 US
; CURRENT APPLICATION NUMBER: US/09/932,613
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-932-613-445

Query Match 100.0%; Score 585; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHREKDLGSENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60

DB 1 DAHSEVAHREKDLGSENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRREV 120

DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRREV 120

QY 121 DVMTAFHNEETFLKKYLIEIARRHPFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180

DB 121 DVMTAFHNEETFLKKYLIEIARRHPFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180

QY 181 KLDELDEGKASSAKQRLKCAASLQKFGERAFAKAWAVARLSQRPFAEFAVSKLVTDLTK 240

DB 181 KLDELDEGKASSAKQRLKCAASLQKFGERAFAKAWAVARLSQRPFAEFAVSKLVTDLTK 240

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 16:07:20 ; Search time 48 seconds
(without alignments)

3359.767 Million cell updates/sec

Title: US-09-832-929-18

Perfect score: 585

Sequence: 1 DAHSEVAHRFKDLEENFK.....TCFAEGKKLVAASQAALGL 585

Scoring table: OLIGO 60.0 , Gapext 60.0

Searched: 1124875 seqs, 275673149 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

- Database : Published Applications AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
 - 6: /cgn2_6/ptodata/2/pubpaa/FCRUS_PUBCOMB.pep:*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
 - 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
 - 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	585	100.0	585	9	US-09-929-552-2	Sequence 2, Appli
2	585	100.0	585	10	US-09-932-613-445	Sequence 445, App
3	585	100.0	585	10	US-09-984-010-26	Sequence 26, Appl
4	585	100.0	585	10	US-09-833-041-18	Sequence 18, Appl
5	585	100.0	585	10	US-09-833-117-18	Sequence 18, Appl
6	585	100.0	585	10	US-09-932-322-445	Sequence 445, App
7	585	100.0	585	11	US-09-832-501-18	Sequence 18, Appl
8	585	100.0	585	11	US-09-833-245-18	Sequence 11, Appl
9	585	100.0	585	12	US-10-424-999-11	Sequence 11, Appl
10	585	100.0	585	12	US-10-425-000-33	Sequence 31, Appl
11	585	100.0	585	13	US-10-153-064-5	Sequence 5, Appli
12	585	100.0	585	14	US-10-153-064-90	Sequence 90, Appl
13	585	100.0	585	14	US-10-319-263-1	Sequence 1, Appli
14	585	100.0	585	14	US-10-319-263-2	Sequence 2, Appli
15	585	100.0	585	14	US-10-153-064-93	Sequence 93, Appl

A:Residues: 1-227 <SHA>
A:Cross-references: GB:M64307; GB:M23660; NID:G213302; PIDN:AAA62808.1; PID:G487648
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:keywords: immunoglobulin

Query Match 80.6%; Score 29; DB 2; Length 227;
Best Local Similarity 71.4%; Pred.No. 1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KADDKET 7
: |||||
Db 161 RVDDKET 167

Search completed: April 19, 2004, 12:02:35
Job time : 2.85319 secs

A>Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: B84846
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-970 <STO>
A/Cross-references: GB:AF002093; NID:G2335108; PIDN:AAC02769.1; GSPDB:GN00139
C/Genetics:
A/Gene: At2g41790
A/Map position: 2
C/Superfamily: insulin-degrading enzyme (IDE)

Query Match 86.1%; Score 31; DB 2; Length 970;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ADDKET 7
|||
Db 501 ADDKET 506

RESULT 11
H72536
A/Title: probable oligopeptide transport system permease protein APE1581 - Aeropyrum pernix (stra
C/Species: Aeropyrum pernix
C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C/Accession: H72536
R/Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yanazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A/Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A/Reference number: A72450; MUID:99310339; PMID:10382966
A/Accession: H72536
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-527 <KAW>
A/Cross-references: DDBJ:AP000062; NID:G5105244; PIDN:BAA80581.1; PID:di044367; PID:G510
A/Experimental source: strain K1
C/Genetics:
A/Gene: APE1581

Query Match 83.3%; Score 30; DB 2; Length 527;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KADDKET 7
|||
Db 90 KADDPET 96

RESULT 12
T42230
A/Title: Af4 protein - mouse (fragment)
N/Alternate names: serine/proline-rich PBL protein
C/Species: Mus musculus (house mouse)
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C/Accession: T42230
R/Baskaran, K.; Efruth, F.; Taborn, G.; Copeland, N.G.; Gilbert, J.; Jenkins, N.A.; Ian
submitted to the EMBL Data Library, July 1997
A/Description: Cloning and developmental expression of the murine homolog of the acute 1
A/Reference number: 222090
A/Accession: T42230
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-1211 <BAS>
A/Cross-references: EMBL:AF013131; NID:G2582018; PID:G2582019; PIDN:AAB82427.1
A/Experimental source: strain CD1
C/Genetics:
A/Gene: Af4
A/Map position: 5

Query Match 83.3%; Score 30; DB 2; Length 1211;
Best Local Similarity 85.7%; Pred. No. 3.5e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KADDKET 7
|||
Db 225 KAQDKET 231

RESULT 13
T42625
A/Title: AP-4 protein - mouse
C/Species: Mus musculus (house mouse)
C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C/Accession: T42625
R/Ishard, P.; Depetris, D.; Mattei, M.G.; Perrier, P.; Djabali, M.
Mamm. Genome 9, 1065-1068, 1998
A/Title: CDNA cloning, expression and chromosomal localization of the murine AP-4 gene
A/Reference number: Z22176; MUID:9909257; PMID:980680
A/Accession: T42625
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-1217 <ISM>
A/Cross-references: EMBL:AF074266; NID:G3328189; PID:G3328190; PIDN:AAD08668.1
A/Experimental source: strain M; thymus

Query Match 83.3%; Score 30; DB 2; Length 1217;
Best Local Similarity 85.7%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KADDKET 7
|||
Db 231 KAQDKET 237

RESULT 14
E33989
A/Title: Ig heavy chain V-5-D-J region - African clawed frog
C/Species: Xenopus laevis (African clawed frog)
C/Date: 23-Mar-1990 #sequence_revision 23-Mar-1990 #text_change 21-Jan-2000
C/Accession: E33989
R/Hsu, E.; Schwager, J.; Alt, F.W.
Proc. Natl. Acad. Sci. U.S.A. 86, 8010-8014, 1989
A/Title: Evolution of immunoglobulin genes: V-H families in the amphibian Xenopus.
A/Reference number: A33989; MUID:90046727; PMID:2510156
A/Accession: E33989
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-117 <HSU>
C/Superfamily: immunoglobulin V region; immunoglobulin homology
C/Keywords: heterotrimer; immunoglobulin
F:15-100/Domain: immunoglobulin homology <IMM>

Query Match 80.6%; Score 29; DB 2; Length 117;
Best Local Similarity 85.7%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KADDKET 7
|||
Db 54 KADDPET 60

RESULT 15
A33937
A/Title: Ig light chain (i301) - horn shark
C/Species: Heterodontus francisci (horn shark)
C/Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 21-Jul-2000
C/Accession: A33937
R/Shablott, M.J.; Litman, G.W.
Proc. Natl. Acad. Sci. U.S.A. 86, 4684-4688, 1989
A/Title: Complete nucleotide sequence of primitive vertebrate immunoglobulin light chai
A/Reference number: A33937; MUID:89282835; PMID:2499889
A/Accession: A33937
A/Status: preliminary
A/Molecule type: mRNA

A;Residues: 1-199 <KUN>
A;Cross-references: GB:Z99106; GB:AL09126; NID:G2632653; PIDN:CAB12281.1; PID:G2632774
A;Experimental source: strain 168
C;Genetics:
A;Gene: rsbx
C;Superfamily: Bacillus subtilis sigma-B activity indirect negative regulator rsbx

Query Match 86.1%; Score 31; DB 2; Length 199;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDDKE 6
| | | | |
DB 32 KADDDKE 37

RESULT 6
B81355
Probable integral membrane protein Cj0826 [imported] - Campylobacter jejuni (strain NCTC)
C;Species: Campylobacter jejuni
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C;Accession: B81355
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Kettle, J.M.; Churcher, C.; Basham, D.; Chilling
W.; Quail, M.; Rajandream, M.A.; VanVliet, A.; Whitehead, S.; Barrell
Nature 403, 665-668, 2000
A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A;Reference number: A81250; MUID:20150912; PMID:10688204
A;Accession: B81355
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-341 <PAR>
A;Cross-references: GB:AL139076; GB:AL111168; NID:G6968128; PIDN:CAB73091.1; PID:G696827
A;Experimental source: serotype O2, strain NCTC 11168
C;Genetics:
A;Gene: Cj0826

Query Match 86.1%; Score 31; DB 2; Length 341;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDDKE 6
| | | | |
DB 179 KADDDKE 184

RESULT 7
D89870
Peptide chain release factor 3 [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: D89870
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: D89870
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-520 <KUR>
A;Cross-references: GB:BA000018; PID:G13700823; PIDN:BA842119.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:
A;Gene: prfC

Query Match 86.1%; Score 31; DB 2; Length 520;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADDKET 7
| | | | |
DB 344 ADDKET 349

RESULT 8
T43845
Chaperonin [validated] - Methanococcus thermolithotrophicus
C;Species: Methanococcus thermolithotrophicus
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C;Accession: T43845
R;Furukani, M.; Iida, T.; Yoshida, T.; Maruyama, T.
J. Biol. Chem. 273, 28399-28407, 1998
A;Title: GroupII chaperonin in a thermophilic methanogen, Methanococcus thermolithotrof
A;Reference number: Z22704; MUID:98447698; PMID:9774467
A;Accession: T43845
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-544 <FUR>
A;Cross-references: EMBL:AB015435; NID:G3776139; PIDN:BAA33889.1; PID:G3776140
A;Experimental source: strain DSM2095
C;Function:
A;Description: promotes the folding of the three-dimensional structure of proteins [val
C;Superfamily: molecular chaperone t-complex-type

Query Match 86.1%; Score 31; DB 2; Length 544;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDDKE 6
| | | | |
DB 145 KADDDKE 150

RESULT 9
A47391
serum albumin precursor - rhesus macaque
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
C;Accession: A47391
R;Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993
A;Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bill:
A;Reference number: A47391; MUID:93211971; PMID:8460152
A;Contents: B/B homozygote
A;Accession: A47391
A;Status: preliminary
A;Molecule type: mRNA; protein
A;Residues: 1-600 <WAT>
A;Cross-references: GB:X90463; NID:G342294; PIDN:AAA36906.1; PID:G342295
A;Experimental source: liver
A;Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBIP:128281)
C;Superfamily: serum albumin; serum albumin repeat homology
F;21-194/Domain: serum albumin repeat homology <SA1>
F;213-386/Domain: serum albumin repeat homology <SA2>
F;405-584/Domain: serum albumin repeat homology <SA3>

Query Match 86.1%; Score 31; DB 2; Length 600;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDDKE 6
| | | | |
DB 576 KADDDKE 581

RESULT 10
B84846
Probable zinc proteinase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 27-Oct-2003
C;Accession: B84846
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, I
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999

RESULT 4
JH0581
NADAD ADP-ribosyltransferase (EC 2.4.2.30) - chicken
N:Alternate names: poly(ADP-ribose) synthase
C:Species: Gallus gallus (chicken)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
C:Accession: JH0581
R:Ritter, M.E.; Garnier, J.M.; Jeltsch, J.M.; Niedergang, C.P.
Gene 102, 157-164, 1991
A:Title: Chicken poly(ADP-ribose) synthetase: complete deduced amino acid sequence and
A:Reference number: JH0581; PMID:91340148; PMID:1840535
A:Accession: JH0581
A:Molecule type: mRNA
A:Residues: 1-1011 <ITT>
A:Cross-references: EMBL:X52630; NID:G63742; PIDN:CAA36917.1; PID:G63743
C:Comment: This protein is a chromatin-bound enzyme.
C:Comment: This enzyme catalyses DNA-dependent post-translational modifications of vari
C:Superfamily: NAD+ ADP-ribosyltransferase
C:Keywords: DNA binding; Glycosyltransferase; NAD; nucleus; pentosyltransferase; zinc i
Query Match 88.9%; Score 32; DB 1; Length 1011;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KADDEK 7
||:||||
DB 188 KADEK 194
RESULT 5
D36131
sigma-B activity indirect negative regulator rsbx - Bacillus subtilis
N:Alternate names: serine phosphatase rsbx
C:Species: Bacillus subtilis
C:Date: 30-Nov-1990 #sequence_revision 30-Nov-1990 #text_change 20-Jun-2000
C:Accession: D36131; C27762; C69702
R:Kaiman, S.; Duncan, M.L.; Thomas, S.M.; Price, C.W.
J. Bacteriol. 172, 5575-5585, 1990
A:Title: Similar organization of the sigB and spoIIA operons encoding alternate sigma i
A:Reference number: A36131; MUID:91008924; PMID:2170324
A:Accession: D36131
A:Molecule type: DNA
A:Residues: 1-199 <KAL>
A:Cross-references: GB:W34995; NID:G143457; PIDN:AAA22714.1; PID:gl43461
R:Duncan, M.L.; Kaiman, S.S.; Thomas, S.M.; Price, C.W.
J. Bacteriol. 169, 771-778, 1987
A:Title: Gene encoding the 37,000-dalton minor sigma factor of Bacillus subtilis RNA pc
A:Reference number: A91861; MUID:87109071; PMID:3027048
A:Accession: C27762
A:Molecule type: DNA
A:Residues: 1-179 <DUN>
A:Cross-references: GB:W34995
R:Kunat, F.; Ogawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.: Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; C
A.: Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrati, E
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, V.; Fuma, S.; Galizzi, A.; Galle
iech, J.; Harwood, C.R.; Henauc, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F
Koster, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinio
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlor
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Ser
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyana
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: C69702
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 0.853186 Seconds
(without alignments)
789.208 Million cell updates/sec

Title: US-09-832-929-18_COPY_560_566

Perfect score: 36

Sequence: 1 KADDKET 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	100.0	598	2 T40676	hypothetical prote
2	36	100.0	609	1 ABHUS	serum albumin prec
3	33	91.7	402	2 T00623	hypothetical prote
4	32	88.9	1011	1 JHC581	NAD ADP-ribosyltra
5	31	86.1	199	2 D36131	sigma-B activity i
6	31	86.1	341	2 B81355	probable integral
7	31	86.1	520	2 D99870	peptide chain rele
8	31	86.1	544	2 T43845	chaperonin [valida
9	31	86.1	600	2 A47391	serum albumin prec
10	31	86.1	970	2 B54846	probable zinc prot
11	30	83.3	527	2 H72536	probable oligopept
12	30	83.3	1211	2 T42230	Ar4 protein - mou
13	30	83.3	1217	2 T42625	Ar-4 protein - mou
14	29	80.6	117	2 A33989	Ig heavy chain V-5
15	29	80.6	227	2 A33937	Ig light chain (I3
16	29	80.6	227	2 B33937	Ig light chain (I5
17	29	80.6	242	2 S06942	hypothetical prote
18	29	80.6	364	2 D72214	phosphate butyrylc
19	29	80.6	377	2 T16784	hypothetical prote
20	29	80.6	431	2 F37318	acyl-CoA thioester
21	29	80.6	454	2 T35380	probable membrane
22	29	80.6	1659	2 T15666	hypothetical prote
23	28	77.8	97	2 AC3259	LSU ribosomal prot
24	28	77.8	141	2 G71079	hypothetical prote
25	28	77.8	150	2 A30494	hypothetical prote
26	28	77.8	187	2 S77483	adenylate kinase (
27	28	77.8	192	2 E69280	iron-dependent rep
28	28	77.8	200	2 T07149	steymin precursor
29	28	77.8	208	2 C42365	flagellar assembly

30 28 77.8 222 2 AD3472
31 28 77.8 264 2 G81252
32 28 77.8 278 1 A40399
33 28 77.8 314 2 T45831
34 28 77.8 336 2 T35093
35 28 77.8 365 1 FC5CR2
36 28 77.8 365 1 A36480
37 28 77.8 365 2 C91099
38 28 77.8 365 2 G85944
39 28 77.8 365 2 AG0871
40 28 77.8 366 2 AB0109
41 28 77.8 381 2 C84289
42 28 77.8 391 1 D69689
43 28 77.8 405 2 E84969
44 28 77.8 488 2 B84197
45 28 77.8 522 2 AB1556

methyated-DNA- [pr
NADH2 dehydrogenas
mannose 6-phosphat
plasmid replicatio
hydroxymethylbiban
translation releas
translation releas
peptide chain rele
peptide chain rele
peptide chain rele
peptide chain rele
dihydropterostate sy
response regulator
flagellar hook pro
anthranilate synth
peptide chain rele

ALIGNMENTS

RESULT 1

T40676
hypothetical protein SPBC776.06c - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 04-Mar-2000
C/Accession: T40676
R/Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Wedler, H.; Kutzner, M.; Wambutt, R.
submitted to the EMBL Data Library, January 1999
A/Reference number: Z21944
A/Accession: T40676
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-598 <LYN>
A/Cross-references: EMBL:AL035263; PIDN:CAA22879.1; GSPDB:GN00067; SPDB:SPBC776.06c
A/Experimental source: strain 972h-; cosmid c776
C/Genetics:
A/Gene: SPDB:SPBC776.06c
A/Map position: 2
C/Supfamily: Schizosaccharomyces pombe hypothetical protein SPBC776.06c

Query Match 100.0%; Score 36; DB 2; Length 598;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KADDKET 7
Db 416 KADDKET 422

RESULT 2

ABHUS
serum albumin precursor [validated] - human
N/Alternate names: preproalbumin
N/Contains: kinetensin
C/Species: Homo sapiens (man)
C/Date: 29-Jul-1981 #sequence_revision 31-Jan-1997 #text_change 17-Mar-2000
C/Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S:
R/Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebt
Nucleic Acids Res. 9, 6103-6114, 1981
A/Title: The sequence of human serum albumin cDNA and its expression in Escherichia col
A/Reference number: A93743; MUID:82081882; PMID:6171778
A/Accession: A93743
A/Molecule type: mRNA
A/Residues: 1-419,'K',421-609 <LAW>
A/Cross-references: EMBL:V00495; GB:J00078; GB:L00132; GB:L00133; NID:g28591; PIDN:CAA:
R/Dugaiczky, A.; Law, S.W.; Dennison, O.E.
Proc Natl. Acad. Sci. U.S.A. 79, 71-75, 1982
A/Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.
A/Reference number: A93936; MUID:82105994; PMID:6275391
A/Accession: A93936
A/Molecule type: mRNA
A/Residues: 1-120,'G',122-609 <DUG>

```
; Sequence 7, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 7
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-7

Query Match      100.0%; Score 45; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TESLVNRRP 9
Db      502 TESLVNRRP 510

RESULT 15
US-09-976-594-977
; Sequence 977, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 977
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No. 6673549 088957CD1
US-09-976-594-977

Query Match      100.0%; Score 45; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TESLVNRRP 9
Db      502 TESLVNRRP 510

Search completed: April 19, 2004, 12:05:21
Job time : 1.3795 secs
```

```

; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5

Query Match      100.0%; Score 45; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
DB 478 TESLVNRRP 486

RESULT 11
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5852352
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri
; APPLICANT: Lyons, David
; APPLICANT: Wurfel, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,619
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-222-619-3

Query Match      100.0%; Score 45; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
DB 502 TESLVNRRP 510

RESULT 12
US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828
; GENERAL INFORMATION:
; APPLICANT: Sreekrishna, Kotikanyadan
; APPLICANT: Barr, Kathryn A.
; APPLICANT: Brierley, Russell A.
; APPLICANT: Thill, Gregory P.
; APPLICANT: Techopp, Juerg F.
; TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/433,037
; FILING DATE: 03-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9108Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-433-037-4

Query Match      100.0%; Score 45; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
DB 502 TESLVNRRP 510

RESULT 13
US-08-897-956A-2
; Sequence 2, Application US/08897956A
; Patent No. 6423512
; GENERAL INFORMATION:
; APPLICANT: Mary Ellen Digan
; APPLICANT: Philip Lake
; APPLICANT: Hermann Gram
; TITLE OF INVENTION: Fusion Polypeptides
; FILE REFERENCE: 600-7244/CPA
; CURRENT APPLICATION NUMBER: US/08/897,956A
; CURRENT FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/022,689
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-08-897-956A-2

Query Match      100.0%; Score 45; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
DB 502 TESLVNRRP 510

RESULT 14
US-10-153-064-7

```

Patent No. 5948609
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C
; APPLICANT: HO, JOSEPH X
; APPLICANT: RUKER, FLORIAN
; TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT
; FILE REFERENCE: 08/984,176
; CURRENT APPLICATION NUMBER: US/08/984,176
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-984-176-1

Query Match 100.0%; Score 45; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.12; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
DB 478 TESLVNRRP 486

RESULT 8
US-08-702-572-2
; Sequence 2, Application US/08702572
; Patent No. 5965386
; GENERAL INFORMATION:
; APPLICANT: Kerry-Williams, Sean M
; APPLICANT: Gilbert, Sarah C
; TITLE OF INVENTION: Yeast Strains and Modified Albumins
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Centeon L.L.C.
; STREET: 1020 First Avenue
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-1310
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,572
; FILING DATE: 11-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 95/23857
; FILING DATE: 1-MAR-1995
; APPLICATION NUMBER: GB 9404270.2
; FILING DATE: 5-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Naomi Biswas
; REGISTRATION NUMBER: 38,384
; REFERENCE/DOCKET NUMBER: CE0114 US
; TELEPHONE: 610/878/4294
; TELEFAX: 610/878/4221
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-702-572-2

Query Match 100.0%; Score 45; DB 2; Length 585;

Best Local Similarity 100.0%; Pred. No. 0.12; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
DB 478 TESLVNRRP 486

RESULT 9
US-08-769-746-2
; Sequence 2, Application US/08769746
; Patent No. 6274305
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; APPLICANT: Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,746
; FILING DATE: 19-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-769-746-2

Query Match 100.0%; Score 45; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.12; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
DB 478 TESLVNRRP 486

RESULT 10
US-10-153-064-5
; Sequence 5, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PFS56
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585

```

; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-448-196A-6

Query Match 100.0%; Score 45; DB 1; Length 583;
Best Local Similarity 100.0%; Pred. No. 0.12; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
Db 477 TESLVNRRP 485

RESULT 5
US-08-153-799-14
; Sequence 14, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Region
; LOCATION: 369..419
; OTHER INFORMATION: /note= "Alternative C-termini of

; OTHER INFORMATION: HSA(1-n)"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..585
; OTHER INFORMATION: /note= "Amino acid sequence of
; natural HSA"
US-08-153-799-14

Query Match 100.0%; Score 45; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.12; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
Db 478 TESLVNRRP 486

RESULT 6
US-08-448-196A-3
; Sequence 3, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
; TITLE OF INVENTION: RELATED PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NASA
; STREET: MARSHALL SPACE FLIGHT CENTER
; CITY: HUNTSVILLE
; STATE: ALABAMA
; COUNTRY: USA
; ZIP: 35812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,196A
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: BROAD JR., ROBERT L.
; REGISTRATION NUMBER: 18,757
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 205-544-0021
; TELEFAX: 205-544-0258
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-448-196A-3

Query Match 100.0%; Score 45; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.12; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
Db 478 TESLVNRRP 486

RESULT 7
US-08-984-176-1
; Sequence 1, Application US/08984176
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:40:29 ; Search time 1.3795 Seconds
(without alignments)
336.813 Million cell updates/sec

Title: US-09-832-929-18_COPY_478_486

Perfect score: 45

Sequence: 1 TESLVNRRP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 segs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA.*

- 1: /cgm2_6/ptodata/2/iaa/5A.COMB.pcp.*
- 2: /cgm2_6/ptodata/2/iaa/5B.COMB.pcp.*
- 3: /cgm2_6/ptodata/2/iaa/6A.COMB.pcp.*
- 4: /cgm2_6/ptodata/2/iaa/6B.COMB.pcp.*
- 5: /cgm2_6/ptodata/2/iaa/PCTUS.COMB.pcp.*
- 6: /cgm2_6/ptodata/2/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	117	1	US-08-448-196A-2
2	45	100.0	582	1	US-08-134-638-1
3	45	100.0	583	1	US-08-448-196A-4
4	45	100.0	583	1	US-08-448-196A-6
5	45	100.0	585	1	US-08-153-739-14
6	45	100.0	585	1	US-08-448-196A-3
7	45	100.0	585	2	US-08-984-176-1
8	45	100.0	585	2	US-08-703-572-2
9	45	100.0	585	3	US-08-769-746-2
10	45	100.0	585	4	US-10-153-064-5
11	45	100.0	609	1	US-08-223-619-3
12	45	100.0	609	1	US-08-433-037-4
13	45	100.0	609	4	US-08-897-956A-2
14	45	100.0	609	4	US-10-153-064-7
15	45	100.0	609	4	US-09-976-594-977
16	45	100.0	609	5	PCT-US95-04075-3
17	45	100.0	610	2	US-08-797-689-2
18	45	100.0	610	4	US-09-984-186-2
19	45	100.0	651	4	US-10-153-064-133
20	45	100.0	652	4	US-10-153-064-96
21	45	100.0	652	4	US-10-153-064-99
22	45	100.0	652	4	US-10-153-064-105
23	45	100.0	652	4	US-10-153-064-132
24	45	100.0	653	4	US-10-153-064-131
25	45	100.0	656	4	US-10-153-064-130
26	45	100.0	660	4	US-10-153-064-90
27	45	100.0	660	4	US-10-153-064-93

28	45	100.0	676	4	US-10-153-064-95	Sequence 95, Appl
29	45	100.0	676	4	US-10-153-064-98	Sequence 98, Appl
30	45	100.0	676	4	US-10-153-064-104	Sequence 104, App
31	45	100.0	676	4	US-10-153-064-127	Sequence 127, App
32	45	100.0	676	4	US-10-153-064-129	Sequence 129, App
33	45	100.0	677	4	US-10-153-064-125	Sequence 125, App
34	45	100.0	680	4	US-10-153-064-123	Sequence 123, App
35	45	100.0	684	4	US-10-153-064-92	Sequence 92, Appl
36	45	100.0	783	1	US-08-256-938-2	Sequence 2, Appl
37	45	100.0	787	1	US-08-256-938-4	Sequence 4, Appl
38	45	100.0	787	2	US-08-797-689-16	Sequence 16, Appl
39	45	100.0	787	4	US-09-984-186-16	Sequence 16, Appl
40	45	100.0	978	4	US-08-897-956A-3	Sequence 3, Appl
41	45	100.0	1184	4	US-10-153-064-89	Sequence 89, Appl
42	39	86.7	658	4	US-10-153-064-102	Sequence 102, App
43	39	86.7	652	4	US-10-153-064-101	Sequence 101, App
44	35	77.8	1274	4	US-09-252-991A-20386	Sequence 20386, A
45	34	75.6	1419	4	US-09-252-991A-31822	Sequence 31822, A

ALIGNMENTS

RESULT 1
US-08-448-196A-2
; Sequence 2, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
; TITLE OF INVENTION: RELATED PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NASA
; STREET: MARSHALL SPACE FLIGHT CENTER
; CITY: HUNTSVILLE
; STATE: ALABAMA
; COUNTRY: USA
; ZIP: 35812

COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,196A
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: BROAD JR, ROBERT L.
; REGISTRATION NUMBER: 18,757
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 205-544-0021
; TELEFAX: 205-544-0258
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-448-196A-2

Query Match 100.0%; Score 45; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. NO. 0.02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
|||||

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: Belongs to the dehydroquininate synthase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP000989; BAB67381.1; -.
CC HAMAP; MF_00110; -; 1.
CC InterPro; IPR002658; DHQ synthase.
CC Pfam; PF01761; DHQ synthase; 1.
CC TIGRFAMs; TIGR01357; aroB; 1.
CC Aromatic amino acid biosynthesis; Lyase; NAD; Complete proteome.
CC SEQUENCE 352 AA; 3913 MW; 88B18080FE057D81 CRC64;
CC -----
CC Query Match 90.6%; Score 29; DB 1; Length 352;
CC Best Local Similarity 71.4%; Pred. No. 99;
CC Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 KADDKET 7
CC |||:|
CC Db 220 KADERET 226
CC -----
CC RESULT 15
CC RL28 BRUWE
CC ID RL28 BRUWE STANDARD; PRT; 97 AA.
CC AC Q8YU6; -----
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE 50S ribosomal protein L28.
CC GN RPB OR BMEI0056 OR BR2015.
CC OS Brucella melitensis, and
CC OS Brucella suis
CC OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
CC OC Brucellaceae; Brucella.
CC OX NCBI_TaxID=29459, 29461;
CC RN [1]
CC SEQUENCE FROM N.A.
CC RC SPECIES=B.melitensis; STRAIN=16M / ATCC 23456 / Biotype 1;
CC RX MEDLINE=20020109; PubMed=11756688;
CC RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Muler C., Los T.,
CC Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
CC Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,
CC Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
CC Haselkorn R., Kypides N., Overbeek R.;
CC RT "The genome sequence of the facultative intracellular pathogen
CC Brucella melitensis";
CC RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
CC RN [2]
CC SEQUENCE FROM N.A.
CC RC SPECIES=B.suis; STRAIN=1330 / Biovar 1;
CC RX MEDLINE=2247741; PubMed=12271122;
CC RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.P.,
CC Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
CC Daugherty S.C., Debay R.T., Durkin A.S., Kolonay J.F., Madupu R.,
CC Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,
CC Riedmiller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
CC Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
CC RT "The Brucella suis genome reveals fundamental similarities between
CC animal and plant pathogens and symbionts";
CC RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
CC RN
CC -!- SIMILARITY: Belongs to the L28P family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB009449; AAL51238.1; -.
CC EMBL; AB014490; AAN30905.1; -.
CC PIR; AC3259; AC3259.
CC TIGR; BR2015; -.
CC HAMAP; MF_00373; -; 1.
CC InterPro; IPR001383; Ribosomal L28.
CC Pfam; PF00830; Ribosomal L28; 1.
CC TIGRFAMs; TIGR00009; L28; 1.
CC Ribosomal protein; Complete proteome.
CC SEQUENCE 97 AA; 10869 MW; A482F8816660DF8C CRC64;
CC -----
CC Query Match 77.8%; Score 28; DB 1; Length 97;
CC Best Local Similarity 83.3%; Pred. No. 42;
CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 KADDKKE 6
CC |||:|
CC Db 71 KSDDKE 76
CC -----
CC Search completed: April 19, 2004, 11:52:58
CC Job time : 1.51062 secs

FT NP_BIND 135 138 GTP (BY SIMILARITY).
SQ SEQUENCE 690 AA; 75829 MW; D1808800578565EA CRC64;
Query Match 86.1%; Score 31; DB 1; Length 690;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KADDKKE 6
DB 301 KADDKKE 306
RESULT 12
ID _SPICI STANDARD; PRT; 339 AA.
AC C31163;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L29.
GN RPMC.
OS Spiroplasma citri.
OC Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;
OC Spiroplasmataceae; Spiroplasma.
OX NCBI_TaxID=2133;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=ATCC 27556 / R8A2;
RX MEDLINE=98406077; PubMed=9733727;
RA Le Dantec L., Castroviejo M., Bove J.M., Saillard C.;
RT "Purification, cloning, and preliminary characterization of a
RT Spiroplasma citri ribosomal protein with DNA binding capacity.";
RL J. Biol. Chem. 273:24379-24386(1998).
CC -!- SIMILARITY: Belongs to the L29P family of ribosomal proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF031160; AAC35872.1; -;
DR HAMAP; MF_00374; fused; 1.
DR InterPro; IPR001854; Ribosomal_L29.
DR Pfam; PF00831; Ribosomal_L29; 1.
DR TIGRFAMs; TIGR00012; L29; 1.
DR PROSITE; PS00579; RIBOSOMAL_L29; FALSE_NEG.
KW Ribosomal protein.
FT DOMAIN 1 74 50S RIBOSOMAL PROTEIN L29.
FT DOMAIN 75 339 UNKNOWN.
SQ SEQUENCE 339 AA; 36556 MW; E7E1A949874E12E8 CRC64;
Query Match 83.3%; Score 30; DB 1; Length 339;
Best Local Similarity 85.7%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 KADDKET 7
DB 326 KADDKET 322
RESULT 13
ID _AF4_MOUSE STANDARD; PRT; 1217 AA.
AC O88573;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE AF-4 protein (Proto-oncogene AF4).
GN MLLT2 OR MLLT2H.
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RP TISSUE=Thymus;
RA Isenard P.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Belongs to the AF4 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF074266; AAD08668.1; -;
DR PIR; T42625; T42625.
DR MGD; MGI:1100819; Mllt2h.
DR InterPro; IPR007797; AF-4.
DR Pfam; PF05110; AF-4; 1.
KW Nuclear protein; Proto-oncogene.
FT DOMAIN 261 266 POLY-PRO.
FT DOMAIN 467 476 POLY-SER.
FT DOMAIN 477 483 POLY-GLU.
FT DOMAIN 841 845 POLY-SER.
FT DOMAIN 871 881 POLY-SER.
SQ SEQUENCE 1217 AA; 131773 MW; BD04563ECF074575 CRC64;
Query Match 83.3%; Score 30; DB 1; Length 1217;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 KADDKET 7
DB 231 KADDKET 237
RESULT 14
ID _SULTO STANDARD; PRT; 352 AA.
AC Q98F96;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 3-dehydroquininate synthase (EC 4.2.3.4).
GN AROB OR ST2272.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=JCM 10545 / 7;
RX MEDLINE=21456156; PubMed=11572479;
RA Kawarayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
CC -!- CATALYTIC ACTIVITY: 3-deoxy-arabino-heptulosonate 7-phosphate = 3-
CC dehydroquininate + phosphate.
CC -!- COFACTOR: NAD and a divalent metal cation (By similarity).
CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC second step.
CC -!- SUBUNIT: Monomer (By similarity).

Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PRO0315; ELONGATNFCT.
DR TIGRFAMs; TIGR00503; pfrc; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR PROSITE; PS00301; EFATOR GTP; 1.
KW protein biosynthesis; GTP-binding; Complete proteome.
FT NP_BIND 17 24 GTP (BY SIMILARITY).
FT NP_BIND 85 89 GTP (BY SIMILARITY).
FT NP_BIND 139 142 GTP (BY SIMILARITY).
SQ SEQUENCE 520 AA; 59571 MW; E24783ECAF72E2F CRC64;
Query Match 86.1%; Score 31; DB 1; Length 520;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ADDKET 7
DB 344 ADDKET 349
RESULT 8
ID RF3 STABP STANDARD; PRT; 520 AA.
AC Q8GPR1;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Peptide chain release factor 3 (RF-3).
GN PRFC OR S50720.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RX PubMed=12950922;
RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
Qian Z.-Q., Mao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
RT "Genome-based analysis of virulence genes in a non-biofilm-forming
Staphylococcus epidermidis strain (ATCC 12228).";
RL Mol. Microbiol. 49:1577-1593(2003)
CC -!- FUNCTION: Increases the formation of ribosomal termination
complexes and stimulates activities of RF-1 and RF-2. It binds
guanine nucleotides and has strong preference for UGA stop codons.
CC It may interact directly with the ribosome. The stimulation of
RF-1 and RF-2 is significantly reduced by GTP and GDP, but not by
GMP (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the GTP-binding elongation factor family.
CC pfrc subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

EMBL; AB016746; AAO04317.1; -.
DR HAMAP; MF_00072; -; 1.
DR InterPro; IPR009022; EFG III V.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR000795; EF_GTPbind.
DR InterPro; IPR004548; Pfrc.
DR InterPro; IPR005225; Small_GTP.
DR InterPro; IPR009000; Translat_factor.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PRO0315; ELONGATNFCT.
DR TIGRFAMs; TIGR00503; pfrc; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.

PROSITE; PS00301; EFATOR GTP; 1.
KW Protein biosynthesis; GTP-binding; Complete proteome.
FT NP_BIND 17 24 GTP (BY SIMILARITY).
FT NP_BIND 85 89 GTP (BY SIMILARITY).
FT NP_BIND 139 142 GTP (BY SIMILARITY).
SQ SEQUENCE 520 AA; 59779 MW; C225A67C469C06E0 CRC64;
Query Match 86.1%; Score 31; DB 1; Length 520;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ADDKET 7
DB 344 ADDKET 349
RESULT 9
ID THS METTL STANDARD; PRT; 544 AA.
AC O93624;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thermosome subunit (Chaperonin subunit).
GN THS.
OS Methanococcus thermolithothrophicus.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanococcaceae; Methanothermococcus.
OX NCBI_TaxID=2186;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=DSM 2095;
RX PubMed=9774467;
RA Futurani M., Iida T., Yoshida T., Maruyama T.;
RT "Group II chaperonin in a thermophilic methanogen, Methanococcus
thermolithothrophicus. Chaperone activity and filament-forming
ability.";
RL J. Biol. Chem. 273:28399-28407(1998).
CC -!- FUNCTION: MOLECULAR CHAPERONE; BINDS UNFOLDED POLYPEPTIDES IN
VITRO, AND HAS A WEAK ATPASE ACTIVITY.
CC -!- SUBUNIT: FORMS AN OLIGOMERIC COMPLEX OF EIGHT-MEMBERED RINGS.
CC -!- SIMILARITY: Belongs to the TCP-1 chaperonin family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

EMBL; AB015435; BAA33889.1; -.
DR PIR; T43845; T43845.
DR HSP; P48424; IASS.
DR InterPro; IPR002194; Chaperonin TCP-1.
DR InterPro; IPR001844; Chaperin Cpn60.
DR InterPro; IPR002423; Cpn60/TCP-1.
DR InterPro; IPR008950; GroEL-ATPase.
DR Pfam; PF00118; cpn60_TCP1; 1.
DR PRINTS; PRO0298; CHAPERONIN60.
DR PRINTS; PRO0304; TCOMPLEXTCP1.
DR PROSITE; PS00750; TCP1_1; 1.
DR PROSITE; PS00751; TCP1_2; 1.
DR PROSITE; PS0095; TCP1_3; 1.
KW Chaperone, ATP-binding.
SQ SEQUENCE 544 AA; 58637 MW; 4128761D958CFCE6 CRC64;
Query Match 86.1%; Score 31; DB 1; Length 544;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KADCKE 6
DB 344 ADDKET 349

DR Genew; HGNC:7639; NAP1L3.
DR MIM; 300117; -
DR GO; GO:0005678; C:chromatin assembly complex; TAS.
DR GO; GO:0006334; P:nucleosome assembly; TAS.
DR InterPro; IPR002164; NAP_family.
DR Pfam; PF00956; NAP; 1.
KW Nuclear protein.
FT DOMAIN 33 50 POLY-SER.
FT DOMAIN 54 70 POLY-SER.
FT DOMAIN 189 194 POLY-GLU.
FT CONFLICT 224 224 A -> P (IN REF. 2).
FT CONFLICT 290 290 R -> Q (IN REF. 2).
FT CONFLICT 406 406 R -> W (IN REF. 2).
FT CONFLICT 490 490 R -> Q (IN REF. 2).
SQ SEQUENCE 506 AA; 57565 MW; 3E679BF2308D0A6B CRC64;

Query Match 86.1%; Score 31; DB 1; Length 506;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KADKKE 6
Db 259 KADKKE 264

RESULT 6
RF3_STAAW
ID RF3_STAAW STANDARD; PRT; 520 AA.
AC Q99V72;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Peptide chain release factor 3 (RF-3).
DR PRFC OR SA0877.
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=159879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11491845;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
Cui L., Oguchi A., Aoki K.I., Nagai Y., Ito T., Kanamori M.,
Takatsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y.,
RA Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA Hiraoka H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
Ogasawara N., Hayashi H., Hiramoto K.,
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus."
RL Lancet 357:1225-1240(2001).
CC -!- FUNCTION: Increases the formation of ribosomal termination
complexes and stimulates activities of RF-1 and RF-2. It binds
guanine nucleotides and has strong preference for UGA stop codons.
CC It may interact directly with the ribosome. The stimulation of RF-
CC 1 and RF-2 is significantly reduced by GTP and GDP, but not by GMP
CC (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the GTP-binding elongation factor family.
CC PrfC subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AP003132; BAB42119.1; -
DR PIR; D89870; D89870.
DR HSPSP; P13551; IELO.
DR HAMAP; MF 00072; -; 1.
DR InterPro; IPR000795; EF_GTPbind.
DR InterPro; IPR009022; EFG_III_V.

DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR004548; PrfC.
DR InterPro; IPR005225; Small_GTP.
DR InterPro; IPR009000; Translat_factor.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR TIGRFAMS; TIGR00503; PrfC; 1.
DR TIGRFAMS; TIGR00231; small_GTP; 1.
DR PROSITE; PS00301; EFACOR_GTP; 1.
KW Protein biosynthesis; GTP-binding; Complete proteome.
FT NP_BIND 17 24 GTP (BY SIMILARITY).
FT NP_BIND 85 89 GTP (BY SIMILARITY).
FT NP_BIND 139 142 GTP (BY SIMILARITY).
FT NP_BIND 520 AA; 59573 MW; 595C9F9D860C4399 CRC64;
SQ SEQUENCE 520 AA; 59573 MW; 595C9F9D860C4399 CRC64;

Query Match 86.1%; Score 31; DB 1; Length 520;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ADDKET 7
Db 344 ADDKET 349

RESULT 7
RF3_STAAW
ID RF3_STAAW STANDARD; PRT; 520 AA.
AC Q8NXC0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Peptide chain release factor 3 (RF-3).
DR PRFC OR MW0901.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=32040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.,
RT "Genome and virulence determinants of high virulence community-
acquired MRSA."
RL Lancet 359:1819-1827(2002).
CC -!- FUNCTION: Increases the formation of ribosomal termination
complexes and stimulates activities of RF-1 and RF-2. It binds
guanine nucleotides and has strong preference for UGA stop codons.
CC It may interact directly with the ribosome. The stimulation of RF-
CC 1 and RF-2 is significantly reduced by GTP and GDP, but not by GMP
CC (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the GTP-binding elongation factor family.
CC PrfC subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AP004825; BAB94766.1; -
DR HAMAP; MF 00072; -; 1.
DR InterPro; IPR000795; EF_GTPbind.
DR InterPro; IPR009022; EFG_III_V.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR004548; PrfC.
DR InterPro; IPR005225; Small_GTP.
DR InterPro; IPR009000; Translat_factor.
DR Pfam; PF00009; GTP_EFTU; 1.

RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Enlian K.D., Errington J., Fabre C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Gallier N.,
RA Gilm S.Y., Glaser P., Gage K., Hagedorn J., Harwood C.R., Henaut A.,
RA Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Jones L.,
RA Hilbert B., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaert-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lepidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medig C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogawa K., Oudega B., Park S.H.,
RA Paro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Prescan B., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche B., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scallan E., Schleich S., Schroeter R., Scoffone F.,
RA Segiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Solido B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpetra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vanderbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zmistein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
RN [4]
RP SEQUENCE OF 1-179 FROM N.A.
RC STRAIN=168;
RX MEDLINE=87109071; PubMed=3027048;
RA Duncan M.L., Kalman S.S., Thomas S.M., Price C.W.;
RT "Gene encoding the 37,000-dalton minor sigma factor of Bacillus
RT subtilis RNA polymerase: isolation, nucleotide sequence, chromosomal
RT locus, and cryptic function.";
RL J. Bacteriol. 169:771-778(1987).
RN [5]
RP FUNCTION.
RC STRAIN=PB2;
RX MEDLINE=92276352; PubMed=1592822;
RA Boylan S.A., Rutherford A., Thomas S.M., Price C.W.;
RT "Activation of Bacillus subtilis transcription factor sigma B by a
RT regulatory pathway responsive to stationary-phase signals.";
RL J. Bacteriol. 174:3695-3706(1992).
RN [6]
RP FUNCTION.
RC STRAIN=PY22;
RX MEDLINE=93224457; PubMed=8468294;
RA Benson A.K., Haldenwang W.G.;
RT "Regulation of sigma B levels and activity in Bacillus subtilis.";
RL J. Bacteriol. 175:2347-2356(1993).
RN [7]
RP FUNCTION.
RC STRAIN=PB2;
RX MEDLINE=96421969; PubMed=8824586;
RA Yang X., Kang C.M., Brody M.S., Price C.W.;
RT "Opposing pairs of serine protein kinases and phosphatases transmit
RT signals of environmental stress to activate a bacterial transcription
RT factor.";
RL Genes Dev. 10:2265-2275(1996).
RN [8]
RP FUNCTION.
RC STRAIN=PY22;
RX MEDLINE=98324978; PubMed=9658013;
RA Smirnova N., Scott J., Voelker U., Haldenwang W.G.;
RT "Isolation and characterization of Bacillus subtilis sigB operon
RT mutations that suppress the loss of the negative regulator RsbX.";
RL J. Bacteriol. 180:3671-3680(1998).
CC -!- FUNCTION: Negative regulator of sigma-B activity. Dephosphorylates
CC rbsB. Plays a role both in maintaining low sigma-B activity during
CC growth and in reestablishing prestress sigma-B activity after
CC induction. Could have a negative feedback role by indirectly
CC communicating sigma-B protein levels.

CC -!- CATALYTIC ACTIVITY: Phosphoserine + H(2)O = serine + phosphate.
CC -!- SIMILARITY: Contains 1 PP2C-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M34995; AAA22714.1; -;
DR EMBL; AB001488; BAA19311.1; -;
DR EMBL; Z99106; CAB12281.1; -;
DR PIR; D36131; D36131.
DR Subtilisin; BG10736; rsbx.
DR InterPro; IPR001932; PP2C-like.
DR SMART; SM00331; PP2C_SIG; 1.
DR SMART; SM00332; PP2C; 1.
KW Hydrolase; Complete proteome. PP2C-LIKE.
FT DOMAIN 22 199
SQ SEQUENCE 199 AA; 22144 MW; 2AAEFB96FB072E33 CRC64;

Query Match 86.1%; Score 31; DB 1; Length 199;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADKKE 6
DB 32 KADKKE 37
|||||

RESULT 5
NPL3_HUMAN STANDARD; PRT; 506 AA.
ID NPL3_HUMAN
AC Q99457; O60788;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 26-FEB-2003 (Rel. 41, Last annotation update)
DE Nucleosome assembly protein 1-like 3.
GN NAP1L3 OR BNAF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=97130622; PubMed=8976385;
RA Watanabe T.K., Fujiwara T., Nakamura Y., Hirai Y., Maekawa H.,
RA Takahashi E.;
RT "Cloning, expression pattern and mapping to Xq of NAP1L3, a gene
RT encoding a peptide homologous to human and yeast nucleosome assembly
RL proteins"; Cell Genet. 74:281-285(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Wilkinson J.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Belongs to the nucleosome assembly protein (NAP)
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D50370; BAA08904.1; -;
DR EMBL; AL009173; -; NOT_ANNOTATED_CDS.

DR PDB; 3PAX; 27-MAY-98.
DR PDB; 4PAX; 27-MAY-98.
DR PDB; 1A26; 27-MAY-98.
DR PDB; 1EFY; 17-JAN-01.
DR InterPro; IPR001357; BRCT.
DR InterPro; IPR008288; NAD_ADPRT.
DR InterPro; IPR001290; PARP.
DR InterPro; IPR004102; PARP_reg.
DR InterPro; IPR008893; WGR.
DR InterPro; IPR001510; Znf_PolyADPpol.
DR Pfam; PF00533; BRCT; 1.
DR Pfam; PF00644; PARP; 1.
DR Pfam; PF02877; PARP_reg; 1.
DR Pfam; PF05406; WGR; 1.
DR Pfam; PF00645; ZF-PARP; 2.
DR PIRSF; PIRSF000489; NAD_ADPRT; 1.
DR ProDom; PD004675; Znf_PolyADPpol; 2.
DR SMART; SM00292; BRCT; 1.
DR PROSITE; PS00172; BRCT; 1.
DR PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR PROSITE; PS0064; PARP_ZN_FINGER_2; 2.
DR TRANSF; PS0064; PARP_ZN_FINGER_2; 2.
KW Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
ADP-ribosylation; Zinc-finger; Zinc; 3D-structure.
FT DNA_BIND 1 370 AUTOMODIFICATION DOMAIN.
FT DOMAIN 371 522
FT DOMAIN 382 458
FT DOMAIN 523 1011
FT ZN_FING 21 56
FT ZN_FING 125 162
FT ZN_FING 207 209
FT DOMAIN 220 225
FT DOMAIN 403 403
FT MOD_RES 404 404
FT MOD_RES 410 410
FT MOD_RES 411 411
FT MOD_RES 432 432
FT MOD_RES 434 434
FT MOD_RES 441 441
FT MOD_RES 442 442
FT MOD_RES 443 443
FT MOD_RES 454 454
FT MOD_RES 468 468
FT MOD_RES 481 481
FT MOD_RES 485 485
FT MOD_RES 488 488
FT MOD_RES 509 509
FT MOD_RES 510 510
FT MOD_RES 517 517
FT CONFLICT 895 895
FT HELIX 664 673
FT TURN 674 674
FT HELIX 676 685
FT TURN 686 687
FT STRAND 688 688
FT TURN 690 692
FT HELIX 695 697
FT HELIX 700 718
FT TURN 719 720
FT HELIX 723 736
FT STRAND 739 739
FT TURN 742 743
FT TURN 749 750
FT HELIX 752 775
FT TURN 776 776
FT HELIX 785 793
FT TURN 794 794
FT STRAND 796 800
FT TURN 803 804
FT HELIX 806 817
FT TURN 818 818
FT HELIX 821 823
FT TURN 824 825
FT STRAND 826 826

FT TURN 839 840
FT HELIX 841 845
FT HELIX 846 850
FT STRAND 854 860
FT HELIX 863 865
FT HELIX 866 872
FT TURN 879 880
FT HELIX 883 885
FT TURN 887 888
FT STRAND 892 895
FT HELIX 898 902
FT HELIX 903 905
FT STRAND 913 922
FT STRAND 926 929
FT TURN 940 941
FT STRAND 944 947
FT STRAND 949 953
FT TURN 955 956
FT STRAND 959 961
FT TURN 962 963
FT STRAND 964 966
FT STRAND 971 973
FT STRAND 981 983
FT STRAND 985 988
FT HELIX 991 993
FT STRAND 994 1006
SQ SEQUENCE 1011 AA; 113520 MW; 261AED9383139144 CRC64;
Query Match 88.9%; Score 32; DB 1; Length 1011;
Best Local Similarity 85.7%; Pred. No. 71;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KADDKET 7
Db 188 KADDKET 194
RESULT 4
ID RSBX BACSU STANDARD; PRT; 199 AA.
AC P17906;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Phosphoserine phosphatase rsbx (EC 3.1.3.3) (Sigma-B negative
DE effector).
GN RSX OR BSU04740.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_taxid=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=91008924; PubMed=2170324;
RA Kalman S., Duncan M.L., Thomas S.M., Price C.W.;
RT "Similar organization of the sigB and spoIIA operons encoding
RT alternate sigma factors of Bacillus subtilis RNA polymerase.";
RL J. Bacteriol. 172:5575-5585(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=91008924; PubMed=2170324;
RA Kalman S., Duncan M.L., Thomas S.M., Price C.W.;
RT "Similar organization of the sigB and spoIIA operons encoding
RT alternate sigma factors of Bacillus subtilis RNA polymerase.";
RL J. Bacteriol. 172:5575-5585(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,


```
Query Match      100.0%; Score 36; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KADDKET 7
Db      584 KADDKET 590

RESULT 2
DDL_BIFLO      STANDARD;      PRT;      428 AA.
AC      Q867C4;
DT      15-MAR-2004 (Rel. 43, Created)
DT      15-MAR-2004 (Rel. 43, Last sequence update)
DE      D-alanine--D-alanine ligase (EC 6.3.2.4) (D-alanylalanine synthetase)
DE      (D-Ala-D-Ala ligase).
GN      DDL OR DDLA OR BL0345.
OS      Bifidobacterium longum.
OC      Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC      Bifidobacteriaceae; Bifidobacterium.
OX      NCBI_TaxID=216816;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=22294977; PubMed=12381787;
RA      Schell M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,
RA      Pessi G., Zwahlen M.-C., Desiere F., Bork P., Delley M.,
RA      Pridmore R.D., Arigoni F.;
RT      "The genome sequence of Bifidobacterium longum reflects its adaptation
RT      to the human gastrointestinal tract.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
CC      -!- FUNCTION: Cell wall formation (by similarity).
CC      -!- CATALYTIC ACTIVITY: ATP + 2 D-alanine = ADP + phosphate + D-
CC      alanyl-D-alanine.
CC      -!- PATHWAY: D-alanine branch of peptidoglycan biosynthesis; second
CC      step.
CC      -!- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC      -!- SIMILARITY: Belongs to the D-alanine--D-alanine ligase family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; AE014653; AAN24184.1; -.
CC      HAMAP; MF_000477; -.
CC      InterPro; IPR000291; Dala_lig_Van.
CC      Pfam; PF01820; Dala_Data_Ligas; 1.
CC      PROSITE; PS00843; DALA_DALA_LIGASE_1; 1.
CC      PROSITE; PS00844; DALA_DALA_LIGASE_2; FALSE NEG.
KW      Ligase; Cell wall; Peptidoglycan synthesis; Complete proteome.
SQ      SEQUENCE 428 AA; 46045 MW; D5017728024C8F46 CRC64;

Query Match      91.7%; Score 33; DB 1; Length 428;
Best Local Similarity 85.7%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 KADDKET 7
Db      262 KADDKET 268

RESULT 3
PPOL_CHICK
ID      PPOL_CHICK      STANDARD;      PRT;      1011 AA.
AC      P26436;
DT      01-AUG-1992 (Rel. 23, Created)
```

```
DT      15-JUL-1998 (Rel. 36, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)-
DE      ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).
GN      ADPRT.
OS      Gallus gallus (Chicken).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Archosauia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC      Gallus.
OX      NCBI_TaxID=9031;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Oviduct;
RX      MEDLINE=91340148; PubMed=1840535;
RA      Ittel M.-E., Garnier J.-M., Jeltsch J.-M., Niedergang C.;
RT      "Chicken poly(ADP-ribose) synthetase: complete deduced amino acid
RT      sequence and comparison with mammalian enzyme sequences.";
RL      Gene 102:157-164(1991).
RN      [2]
RP      X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 659-1011.
RX      MEDLINE=96353841; PubMed=8755499;
RA      Ruf A., Menissier de Murcia J., de Murcia G.M., Schulz G.E.;
RT      "Structure of the catalytic fragment of poly(AD-ribose) polymerase
RT      from chicken.";
RL      Proc. Natl. Acad. Sci. U.S.A. 93:7481-7485(1996).
RN      [3]
RP      X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 659-1011, AND REVISION TO
RP      895.
RX      MEDLINE=98191351; PubMed=9521710;
RA      Ruf A., de Murcia G.M., Schulz G.E.;
RT      "Inhibitor and NAD+ binding to poly(ADP-ribose) polymerase as derived
RT      from crystal structures and homology modeling.";
RL      Biochemistry 37:3893-3900(1998).
RN      [4]
RP      X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 659-1011.
RX      MEDLINE=98239715; PubMed=9571033;
RA      Ruf A., Rolli V., de Murcia G.M., Schulz G.E.;
RT      "The mechanism of the elongation and branching reaction of poly(ADP-
RT      ribose) polymerase as derived from crystal structures and
RT      mutagenesis.";
RL      J. Mol. Biol. 278:57-65(1998).
CC      -!- FUNCTION: Poly[ADP-ribose] polymerase modifies various nuclear
CC      proteins by poly(ADP-ribose)ation. The modification is dependent
CC      on DNA and is involved in the regulation of various important
CC      cellular processes such as differentiation, proliferation, and
CC      tumor transformation and also in the regulation of the molecular
CC      events involved in the recovery of cell from DNA damage.
CC      -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-riboseyl](N)-acceptor =
CC      nicotinamide + [ADP-D-riboseyl](N+1)-acceptor.
CC      -!- COFACTOR: Zinc. Contains two moles of zinc per mole of protein.
CC      -!- SUBUNIT: Homodimer (Potential).
CC      -!- SUBCELLULAR LOCATION: Nuclear.
CC      -!- MISCELLANEOUS: The ADP-D-riboseyl group of NAD(+) is transferred to
CC      an acceptor carboxyl group on a histone or the enzyme itself, and
CC      further ADP-riboseyl groups are transferred to the 2'-position of
CC      the terminal adenosine moiety, building up a polymer with an
CC      average chain length of 20-30 units.
CC      -!- SIMILARITY: Belongs to the PARP family.
CC      -!- SIMILARITY: Contains 1 BRCT domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; X52690; CAA36917.1; -.
CC      PIR; JH0581; JH0581.
CC      PDB; 2PAM; 27-MAY-98.
CC      PDB; 1PAX; 15-MAY-97.
CC      PDB; 2PAX; 27-MAY-98.
```

RC TISSUE=Liver, and Skeletal muscle;
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketting A.M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [8]
RN SEQUENCE OF 25-609.
RX MEDLINE=76187907; PubMed=1225573;
RA Meloun B., Moravsek L., Kostka V.;
RT "Complete amino acid sequence of human serum albumin";
RL FEBS Lett. 58:134-137(1975).
RN [9]
RN SEQUENCE OF 25-609.
RX Brown J.R., Shockley P., Behrens P.O.;
RA (In) Bing D.H. (eds.);
RL The chemistry and physiology of the human plasma proteins, pp.23-40,
RL Pergamon Press, New York (1979).
RN [10]
RN SEQUENCE OF 1-455 FROM N.A.
RC TISSUE=Liver;
RA Menaya J., Parrilla R., Ayuso M.S.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [11]
RN SEQUENCE OF 1-26 FROM N.A.
RX MEDLINE=86140099; PubMed=2419329;
RA Urao Y., Watanabe K., Sakai M., Tamaoki T.;
RT "The human albumin gene. Characterization of the 5' and 3' flanking
RT regions and the polymorphic gene transcripts";
RL J. Biol. Chem. 261:3244-3251(1986).
RN [12]
RN SEQUENCE OF 222-229.
RX MEDLINE=76257808; PubMed=955075;
RA Walker J.E.;
RT "Lysine residue 199 of human serum albumin is modified by
RT acetylsalicylic acid";
RL FEBS Lett. 66:173-175(1976).
RN [13]
RN SEQUENCE OF 25-44 AND 480-499.
RC TISSUE=Heart;
RX MEDLINE=95203287; PubMed=7895732;
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
RT "The human myocardial two-dimensional gel protein database: update
RT 1994";
RL Electrophoresis 15:1459-1465(1994).
RN [14]
RN DISULFIDE BONDS.
RA Sabar M.A., Stockbauer P., Moravsek L., Meloun B.;
RT "Disulfide bonds in human serum albumin";
RL Collect. Czech. Chem. Commun. 42:564-579(1977).
RN [15]
RN BILIRUBIN-BINDING SITE.
RX MEDLINE=78186630; PubMed=656055;
RA Jacobsen C.;
RT "Lysine residue 240 of human serum albumin is involved in high-
RT affinity binding of bilirubin";
RL Biochem. J. 171:453-459(1978).
RN [16]
RN VARIANT CANTERBURY ASN-337.
RX MEDLINE=87157744; PubMed=3828358;
RA Brennan S.O., Herbert P.;
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second
RT domain of serum albumin";
RL Biochim. Biophys. Acta 912:191-197(1987).
RN [17]
RN VARIANTS NAG-2 AND NAG-3.
RX MEDLINE=88068523; PubMed=3479777;
RA Takashashi N., Takanashi Y., Isobe T., Putnam F.W., Fujita M.,
RA Satoh C., Neel J.V.;
RT "Amino acid substitutions in inherited albumin variants from
RT Amerindian and Japanese populations";
RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).
RN [18]
RN VARIANTS NAG-1, HIR-1, HIR-2 AND TOCHIGI.
RX MEDLINE=89345611; PubMed=2762316;
RA Arai K., Madison J., Huse K., Ishioh N., Satoh C., Fujita M.,
RA Neel J.V., Sakurabayashi I., Putnam F.W.;
RT "Point substitutions in Japanese allolbumins";
RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).
RN [19]
RN VARIANTS MANAUS, OSAKA, NAGOYA, FUKUOKA, HONOLULU AND NEW-GUINEA.
RX MEDLINE=90115905; PubMed=2404284;
RA Arai K., Madison J., Shimizu A., Putnam F.W.;
RT "Point substitutions in albumin genetic variants from Asia";
RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).
RN [20]
RN DESCRIPTION OF VARIANT REDHILL.
RX MEDLINE=90115852; PubMed=2104980;
RA Brennan S.O., Wyles T., Peach R.J., Donaldson D., George P.M.;
RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of
RT human serum albumin whose precursor has an aberrant signal peptidase
RT cleavage site";
RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).
RN [21]
RN VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.
RX MEDLINE=91062352; PubMed=2247440;
RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,
RA Watkins S., Putnam F.W.;
RT "Mutations in genetic variants of human serum albumin found in
RT Italy";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).
RN [22]
RN VARIANT VENEZIA.
RX MEDLINE=91296740; PubMed=2068071;
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,
RA Minchiotti L., Putnam F.W.;
RT "A donor splice mutation and a single-base deletion produce two
RT carboxyl-terminal variants of human serum albumin";
RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).
RN [23]
RN VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;
RX KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.
RX MEDLINE=92052189; PubMed=1946412;
RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,
RA Matsuda Y.-I., Amaki I., Putnam F.W.;
RT "Genetic variants of serum albumin in Americans and Japanese";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).
RN [24]
RN VARIANT CASEBROOK ASN-518.
RX MEDLINE=91316157; PubMed=1859851;
RA Peach R.J., Brennan S.O.;
RT "Structural characterization of a glycoprotein variant of human serum
RT albumin: albumin Casebrook (494 Asp-->Asn)";
RL Biochim. Biophys. Acta 1097:49-54(1991).
RN [25]
RN VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.
RX MEDLINE=92190239; PubMed=1347703;
RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,
RA Rochu D., Porta F.;
RT "Two allolbumins with identical electrophoretic mobility are produced

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 0.510619 Seconds

Title: Perfect score: 36 (without alignments)

Sequence: 1 KADDKET 7 713.823 Million cell updates/sec

Scoring table: BLOSUM62

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	609	1	ALBU HUMAN
2	33	91.7	428	1	DBL BIFLO
3	32	89.9	1011	1	PPOL CHICK
4	31	86.1	199	1	RSBX_BACSU
5	31	86.1	506	1	NPL3_HUMAN
6	31	86.1	520	1	REF3_STAAN
7	31	86.1	520	1	REF3_STAAN
8	31	86.1	520	1	REF3_STAAN
9	31	86.1	544	1	THS METTL
10	31	86.1	600	1	ALBU MACMU
11	31	86.1	690	1	REG BRAJA
12	30	83.3	339	1	RL23 SPICI
13	30	83.3	1217	1	AF4 MOUSE
14	29	80.6	352	1	AF0B SULTO
15	28	77.8	97	1	RL28 BRUME
16	28	77.8	185	1	KAD1 SYNY3
17	28	77.8	200	1	SVST LYCES
18	28	77.8	208	1	FLIH_BACSU
19	28	77.8	266	1	ZNIN_RAT
20	28	77.8	278	1	MRPD MOUSE
21	28	77.8	316	1	HEM3 PROMP
22	28	77.8	336	1	HEM3 SCHPO
23	28	77.8	365	1	RF2 ECOL6
24	28	77.8	365	1	RF2 ECOLI
25	28	77.8	365	1	RF2 ECOLI
26	28	77.8	366	1	RF2 YERPE
27	28	77.8	391	1	RF2 YERPE
28	28	77.8	405	1	FLGE BUCAI
29	28	77.8	497	1	CBS DICDI
30	28	77.8	511	1	Y8B7 BACTIN
31	28	77.8	522	1	RF3 LISIN
32	28	77.8	522	1	RF3 LISMO
33	28	77.8	556	1	SVRI_BACAA

ALIGNMENTS

RESULT 1

ID	ALBU HUMAN	STANDARD;	PRT;	609 AA.
AC	P02768; Q95574; Q13140; Q9P157; Q9P117; Q9UHS3; Q9UJZ0;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Serum albumin precursor.			
GN	ALB.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=86196112; PubMed=3009475;			
RA	Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,			
RA	Beattie W.G., Dugaiczky A.;			
RT	"Molecular structure of the human albumin gene is revealed by			
RT	nucleotide sequence within q11-22 of chromosome 4.";			
RL	J. Biol. Chem. 261:6747-6757(1986).			
RN	[2]			
RP	SEQUENCE FROM N.A., AND VARIANT LYS-420.			
RA	MEDLINE=82081882; PubMed=6171778;			
RA	Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,			
RA	Najarian R.C., Seeburg P.H., Wion K.L.;			
RT	"The sequence of human serum albumin cDNA and its expression in E.			
RT	coli.";			
RL	Nucleic Acids Res. 9:6103-6114(1981).			
RN	[3]			
RP	SEQUENCE FROM N.A., AND VARIANT GLY-121.			
RA	MEDLINE=82105994; PubMed=6275391;			
RA	Dugaiczky A., Law S.W., Dennison O.E.;			
RT	"Nucleotide sequence and the encoded amino acids of human serum			
RT	albumin mRNA.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	TISSUE=Liver;			
RA	Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;			
RT	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).			
RA	TISSUE=fetal liver;			
RA	Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,			
RA	Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;			
RT	"Functional prediction of the coding sequences of 121 new genes			
RT	deduced by analysis of cDNA clones from human fetal liver.";			
RL	Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.			
RA	Huang M.C., Wu H.T.;			
RT	"The cDNA sequences of human serum albumin.";			
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.			
RN	[7]			
RP	SEQUENCE FROM N.A.			

Q814q8 bacillus ce
P42971 bacillus su
P35209 saccharomyc
P15711 theileria p
P19838 homo sapien
Q12451 saccharomyc
P19296 schizosacch
Q92910 bacillus ha
Q9cdw6 lactococcus
Q99s25 staphylococ
Q8crg4 staphylococ
Q9a1x0 streptococc

Query Match 86.1%; Score 31; DB 4; Length 506;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKKE 6
 DB 259 KADDKKE 264

RESULT 13
 ID Q8CPR1 PRELIMINARY; PRT; 520 AA.
 AC Q8CPR1;
 DT 01-MAR-2003 (TReMBLrel. 23, Created)
 DT 01-MAR-2003 (TReMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Peptide chain release factor RF-3.
 GN SE0720.
 OS Staphylococcus epidermidis.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=1282;
 RN [1]

SEQUENCE FROM N.A.
 STRAIN=ATCC 12228;
 RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,
 Chen Z., Wen Y.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE016746; AAC04317.1; -
 DR GO; GO:0005737; C:cytoplasm; IEA.
 DR GO; GO:0005525; F:GTP binding; IEA.
 DR GO; GO:0003746; F:translation elongation factor activity; IEA.
 DR GO; GO:0016149; F:translation release factor activity, codon . . .; IEA.
 DR GO; GO:0006414; P:translational elongation; IEA.
 DR GO; GO:0006415; P:translational termination; IEA.
 DR InterPro; IPR003022; EFG_III_V.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR000795; EFGPBbind.
 DR InterPro; IPR004548; Pfc.
 DR InterPro; IPR005225; Small_GTP.
 DR InterPro; IPR003000; Translat_factor.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR PRINTS; PR00315; ELONGATNFCT.
 DR TIGRPFAMs; TIGR00503; pfic; 1.
 DR TIGRPFAMs; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00301; EFACITOR_GTP; 1.
 KW Complete proteome.
 SQ SEQUENCE 520 AA; 59779 MW; C225A67C469C06E0 CRC64;

Query Match 86.1%; Score 31; DB 16; Length 520;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADDKET 7
 DB 344 ADDKET 349

RESULT 14
 ID Q9GRW3 PRELIMINARY; PRT; 620 AA.
 AC Q9GRW3;
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Intermediate filament protein.
 OS Glottidia pyramidata.
 OC Eukaryota; Metazoa; Brachiopoda; Linguliformes; Lingulata; Lingulida;
 OC Linguloidae; Lingulidae; Glottidia.
 OX NCBI_TaxID=34515;
 RN [1]

SEQUENCE FROM N.A.

RA Erber A., Riemer D., Bovenschulte M., Weber K.;
 RT "Molecular phylogeny of metazoan intermediate filament proteins.";
 CC Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 DR EMBL; AJ005004; CAC12741.1; -
 DR GO; GO:0005882; C:intermediate filament; IEA.
 DR GO; GO:0005871; C:kinesin complex; IEA.
 DR InterPro; IPR001664; IF.
 DR InterPro; IPR001322; IF_tail_C.
 DR Pfam; PF00038; filament; 1.
 DR PROSITE; PS00932; IF_tail; 1.
 DR PROSITE; PS00226; IF; 1.
 KW Coiled coil; Intermediate filament.
 SQ SEQUENCE 620 AA; 69508 MW; 646D394AEC9151AB CRC64;

Query Match 86.1%; Score 31; DB 5; Length 620;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKKE 6
 DB 549 KADDKKE 554

RESULT 15
 ID Q96NK7 PRELIMINARY; PRT; 632 AA.
 AC Q96NK7;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE Hypothetical protein FLJ30678.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

SEQUENCE FROM N.A.
 RP TISSUE=Brain;
 RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
 RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,
 RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
 RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
 RA Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
 RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
 RA Masuho Y., Nagai K., Isogai T.;
 RA "NEDO human cDNA sequencing project.";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK055240; BAB70886.1; -
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00071; HOMEBOX_2; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 632 AA; 70224 MW; C1ACD1B97B5F95C2 CRC64;

Query Match 86.1%; Score 31; DB 4; Length 632;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKKE 6
 DB 583 KADDKKE 588

Search completed: April 19, 2004, 12:00:18
 Job time : 3.70175 secs

Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADCKE 6
DB 179 KADCKE 184

RESULT 9
Q931U9 ID Q931U9 PRELIMINARY; PRT; 405 AA.
AC Q931U9;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Peptide chain release factor 3.
GN SAVI021.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hoshoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekinizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Shiba T.,
RA Kanesaka M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus".
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003361; BAB57183.1; --
DR GO; GO:0005525; P:GTP binding; IEA.
DR GO; GO:0003746; P:translational elongation factor activity; IEA.
DR GO; GO:0006414; P:translational elongation; IEA.
DR InterPro; IPR009022; EFG III_V.
DR InterPro; IPR004161; EFTU D2.
DR InterPro; IPR000795; EF_GTPbind.
DR InterPro; IPR009000; Translat_factor.
DR Pfam; PF00009; GTP_EFTU_1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
KW Complete proteome.
SQ SEQUENCE 405 AA; 46773 MW; BED0F01974E72BC4 CRC64;

Query Match 86.1%; Score 31; DB 16; Length 405;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADDKET 7
DB 229 ADDKET 234

RESULT 10
Q7X231 ID Q7X231 PRELIMINARY; PRT; 441 AA.
AC Q7X231;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Peptide chain release factor 3 (Fragment).
GN PRFC.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP SEQUENCE FROM N.A.
RX Wootton M., Avison M.B., Bennett P.M., Howe R.A., MacGowan A.P.,
RA Walsh T.R.;
RT "Genetic analysis of seventeen genes in Staphylococcus aureus with
RT reduced susceptibility to vancomycin (VRSA) and hetero-VRSA (hVRSA).";
DR EMBL; AJ564051; CAD91756.1; --
DR GO; GO:0003746; P:translational elongation factor activity; IEA.
DR GO; GO:0006414; P:translational elongation; IEA.
DR InterPro; IPR009022; EFG III_V.
DR InterPro; IPR004161; EFTU D2.
DR InterPro; IPR000795; EF_GTPbind.
DR InterPro; IPR009000; Translat_factor.
DR Pfam; PF00009; GTP_EFTU_1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
KW Complete proteome.
SQ SEQUENCE 405 AA; 46773 MW; BED0F01974E72BC4 CRC64;

Query Match 86.1%; Score 31; DB 16; Length 405;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADDKET 7
DB 229 ADDKET 234

RESULT 10
Q7X231 ID Q7X231 PRELIMINARY; PRT; 441 AA.
AC Q7X231;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Peptide chain release factor 3 (Fragment).
GN PRFC.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP SEQUENCE FROM N.A.
RX Wootton M., Avison M.B., Bennett P.M., Howe R.A., MacGowan A.P.,
RA Walsh T.R.;
RT "Genetic analysis of seventeen genes in Staphylococcus aureus with
RT reduced susceptibility to vancomycin (VRSA) and hetero-VRSA (hVRSA).";
DR EMBL; AJ564051; CAD91756.1; --
DR GO; GO:0003746; P:translational elongation factor activity; IEA.
DR GO; GO:0006414; P:translational elongation; IEA.
DR InterPro; IPR009022; EFG III_V.
DR InterPro; IPR004161; EFTU D2.
DR InterPro; IPR000795; EF_GTPbind.
DR InterPro; IPR009000; Translat_factor.
DR Pfam; PF00009; GTP_EFTU_1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
KW Complete proteome.
SQ SEQUENCE 405 AA; 46773 MW; BED0F01974E72BC4 CRC64;

Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ564050; CAD91755.1; --
FT NON_TER 1
FT NON_TER 441
SQ SEQUENCE 441 AA; 49981 MW; 64D313C80BE025B3 CRC64;

Query Match 86.1%; Score 31; DB 2; Length 441;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADDKET 7
DB 335 ADDKET 340

RESULT 11
Q7X230 ID Q7X230 PRELIMINARY; PRT; 441 AA.
AC Q7X230;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Peptide chain release factor 3 (Fragment).
GN PRFC.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu3;
RA Wootton M., Avison M.B., Bennett P.M., Howe R.A., MacGowan A.P.,
RA Walsh T.R.;
RT "Genetic analysis of seventeen genes in Staphylococcus aureus with
RT reduced susceptibility to vancomycin (VRSA) and hetero-VRSA (hVRSA).";
DR EMBL; AJ564051; CAD91756.1; --
FT NON_TER 1
FT NON_TER 441
SQ SEQUENCE 441 AA; 49981 MW; 64D313C80BE025B3 CRC64;

Query Match 86.1%; Score 31; DB 2; Length 441;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADDKET 7
DB 335 ADDKET 340

RESULT 12
Q8IYV1 ID Q8IYV1 PRELIMINARY; PRT; 506 AA.
AC Q8IYV1;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Nucleosome assembly protein 1-like 3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC034954; AAH34954.1; --
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006334; P:nucleosome assembly; IEA.
DR InterPro; IPR002164; NAP_family.
DR Pfam; PF00956; NAP; 1.
SQ SEQUENCE 506 AA; 57570 MW; B3977473F97C804D CRC64;

QY	1 KADDKET 7
DB	262 KADDRET 268
RESULT 6	
ID	Q95YMO PRELIMINARY; PRT; 673 AA.
AC	Q95YMO;
DT	01-DEC-2001 (TReMBLrel. 19, Created)
DT	01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE	01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE	Pm-integrin alpha (Fragment).
PM-INT.	
GN	Polyandrocampa misakiensis.
OS	Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC	Stolidobranchia; Styelidae; Polyandrocarpa.
NCBI_TaxID=7723;	
SEQUENCE FROM N.A.	(1)
RN	MEDLINE=21430864; PubMed=11546750;
RP	Matsumoto J., Nakamoto C., Fujiwara S., Yubisui T., Kawamura K.;
RA	"A novel C-type lectin regulating cell growth, cell adhesion and cell
RT	differentiation of the multipotent epithelium in budding tunicates.";
RRL	Development 128:3339-3347(2001).
DR	EMBL; AB056090; BAB63365.1; -
GO	GO:0008305; C:integrin complex; IEA.
GO	GO:0004895; F:cell adhesion receptor activity; IEA.
GO	GO:0007160; P:cell-matrix adhesion; IEA.
GO	GO:0007229; P:integrin-mediated signaling pathway; IEA.
InterPro	IPIR000413; Integrin_alpha.
Pfam	PF01839; FG-GAP; 1.
PRINTS	PR01185; INTEGRIN.
PROSITE	PS00242; INTEGRIN_ALPHA; 1.
NON TER	1
FT	SEQUENCE 673 AA; 74761 MW; DIDAA62D2FF9FE34 CRC64;
Query Match	91.7%; Score 33; DB 5; Length 673;
Best Local Similarity	85.7%; Pred No. 1.9e+02;
Matches	6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY	1 KADDKET 7
DB	665 KADDRET 671
RESULT 7	
ID	Q7TOR9 PRELIMINARY; PRT; 281 AA.
AC	Q7TOR9;
DT	01-OCT-2003 (TReMBLrel. 25, Created)
DT	01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DE	01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE	Hypothetical protein.
OS	Xenopus laevis (African clawed frog).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC	Xenopodinae; Xenopus.
NCBI_TaxID=8355;	
[1]	
RN	SEQUENCE FROM N.A.
RP	TISSEU-Spleen;
RC	MEDLINE=22341132; PubMed=12454917;
RX	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA	Richardson P.;
RA	"Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT	initiative";
RL	Dev. Dyn. 225:384-391(2002).
DEV	
SEQUENCE FROM N.A.	(2)
RN	SEQUENCE 281 AA; 30274 MW; F05C3724FAED5DDDB CRC64;
Query Match	86.1%; Score 31; DB 13; Length 281;
Best Local Similarity	100.0%; Pred. No. 2.1e+02;
Matches	6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 KADDKE 6
DB	41 KADDKE 46
RESULT 8	
ID	Q9P97 PRELIMINARY; PRT; 341 AA.
AC	Q9P97;
DT	01-OCT-2000 (TReMBLrel. 15, Created)
DT	01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT	01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE	Putative integral membrane protein.
GN	CJ0826.
OS	Campylobacter jejuni.
OC	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC	Campylobacteraceae; Campylobacter.
NCBI_TaxID=197;	
[1]	
RN	SEQUENCE FROM N.A.
RP	STEAIN-NCTC 11168;
RC	MEDLINE=20150912; PubMed=10688204;
RX	Farkhill J., Wren B.W., Mungall K., Kesley J.M., Churcher C.,
RA	Basham D., Chillingworth T., Davies R.M., Felwell T., Holtroyd S.,
RA	Jagals K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA	Quail M.A., Rajadream M.A., Rutherford K.M., van Vliet A.H.M.,
RA	Whitehead S., Barrell B.G.;
RT	"The genome sequence of the food-borne pathogen Campylobacter jejuni
RT	reveals hypervariable sequences.";
RL	Nature 403:665-668(2000).
DR	EMBL; AL139076; CAB73091.1; -
DR	PIR; B81355; B81355.
DR	GO:0016021; C:integral to membrane; IEA.
DR	InterPro; IPR005495; YjGP_YjGQ.
DR	Pfam; PF03739; YjGP_YjGQ; 1.
KW	Complete proteome.
SEQUENCE	341 AA; 39619 MW; 254C2D8E04489FF1 CRC64;
Query Match	86.1%; Score 31; DB 16; Length 341;
Best Local Similarity	100.0%; Pred. No. 2.1e+02;
Matches	6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 KADDKE 6
DB	41 KADDKE 46
RESULT 9	
ID	Q7TOR9 PRELIMINARY; PRT; 281 AA.
AC	Q7TOR9;
DT	01-OCT-2003 (TReMBLrel. 25, Created)
DT	01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DE	01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE	Hypothetical protein.
OS	Xenopus laevis (African clawed frog).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC	Xenopodinae; Xenopus.
NCBI_TaxID=8355;	
[1]	
RN	SEQUENCE FROM N.A.
RP	TISSEU-Spleen;
RC	MEDLINE=22341132; PubMed=12454917;
RX	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA	Richardson P.;
RA	"Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT	initiative";
RL	Dev. Dyn. 225:384-391(2002).
DEV	
SEQUENCE FROM N.A.	(2)
RN	SEQUENCE 281 AA; 30274 MW; F05C3724FAED5DDDB CRC64;
Query Match	86.1%; Score 31; DB 13; Length 281;
Best Local Similarity	100.0%; Pred. No. 2.1e+02;
Matches	6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 KADDKE 6
DB	41 KADDKE 46
RESULT 10	
ID	Q9P97 PRELIMINARY; PRT; 341 AA.
AC	Q9P97;
DT	01-OCT-2000 (TReMBLrel. 15, Created)
DT	01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT	01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE	Putative integral membrane protein.
GN	CJ0826.
OS	Campylobacter jejuni.
OC	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC	Campylobacteraceae; Campylobacter.
NCBI_TaxID=1	

ID Q86YGO PRELIMINARY; PRT; 417 AA.
 AC Q86YGO;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Similar to alpha-fetoprotein.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OC NCBI_TaxID=9606;
 OX [1]
 RN NCBI_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Strausberg R.;
 RL Submitted (DSC-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL; BC041789; AAH41789.1; -
 DR GO; GO:0005615; C:extracellular space; IEA.
 DR GO; GO:0005386; F:carrier activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000264; Serum albumin.
 DR Pfam; PF00273; transport_P707; 2.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum albumin; 1.
 DR SMART; SM00103; ALBUMIN; 2.
 DR PROSITE; PS00212; ALBUMIN; 2.
 DR PROSITE; PS00212; ALBUMIN; 2.
 SQ SEQUENCE 417 AA; 47360 MW; 16E764833EEF4E8D CRC64;

Query Match 100.0%; Score 36; DB 4; Length 417;
 Best Local Similarity 100.0%; Pred. No. 28; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
 DB 392 KADDKET 398

RESULT 3
 O94674 PRELIMINARY; PRT; 598 AA.
 ID O94674;
 AC O94674;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein.
 DE Hypothetical protein.
 GN SPBC776.06C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OC NCBI_TaxID=4896;
 OX [1]
 RN NCBI_TaxID=4896;
 RP SEQUENCE FROM N.A.
 RC STRAIN=972h-;
 RA Lyne M., Rajandream M.A., Barrell B.G., Wedler H., Kutzner M.,
 RA Wambutt R.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AL035263; CAA22879.1; -
 DR PIR; T40676; T40676.
 DR GeneB_Spombe; SPBC776.06C; -
 DR Hypothetical protein.
 KW Hypothetical protein.
 SQ SEQUENCE 598 AA; 68176 MW; F99AAEA874351D40 CRC64;

Query Match 100.0%; Score 36; DB 3; Length 598;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
 DB 416 KADDKET 422

RESULT 4
 O80595 PRELIMINARY; PRT; 402 AA.
 ID O80595;
 AC O80595;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE T2711.6 protein.
 GN T2711.6.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 OX [1]
 RN NCBI_TaxID=3702;
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Federici N.A., Palm C.J., Conway A.B., Kurtz D.B., Conway A.R.,
 RA Au M., Araujo R., Buehler E., Dewar K., Peng J., Kim C., Li Y.,
 RA Oji O., Osborne B.I., Shinn P., Sun H., Toriumi M., Vysotskaia V.S.,
 RA Yu G., Ecker J., Theologis A., Davis R.W.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC004122; AAC34332.1; -
 DR PIR; T00623; T00623.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR InterPro; IPR008262; Lipase AS.
 DR InterPro; IPR000379; Ser esters.
 DR PROSITE; PS00120; LIPASE_SER; 1.
 DR PROSITE; PS00120; LIPASE_SER; 1.
 SQ SEQUENCE 402 AA; 45654 MW; 8FEC766F3F3C5A24 CRC64;

Query Match 91.7%; Score 33; DB 10; Length 402;
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
 DB 38 KSDDKET 44

RESULT 5
 Q8G7C4 PRELIMINARY; PRT; 428 AA.
 ID Q8G7C4;
 AC Q8G7C4;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE D-alanine-D-alanine ligase.
 GN DDLA OR BL0345
 OS Bifidobacterium longum
 OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
 OC Bifidobacteriaceae; Bifidobacterium.
 OC NCBI_TaxID=216816;
 OX [1]
 RN NCBI_TaxID=216816;
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCC 2705;
 RA MEDLINE=22294977; PubMed=12381787;
 RA Schell M.A., Karimantrou M., Snel B., Vilanova D., Berger B.,
 RA Fessi G., Zwahlen M.-C., Desiere F., Bork P., Delley M.,
 RA Fridmore R.D., Arigoni F.;
 RT "The genome sequence of Bifidobacterium longum reflects its adaptation
 to the human gastrointestinal tract."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
 DR EMBL; AE014653; RAN24184.1; -
 DR GO; GO:0008716; F:D-alanine-D-alanine ligase activity; IEA.
 DR GO; GO:0016874; F:ligase activity; IEA.
 DR GO; GO:0009252; P:peptidoglycan biosynthesis; IEA.
 DR InterPro; IPR000291; Dala_lig_Van.
 DR Pfam; PF01820; Dala_Dala_ligas; 1.
 DR PROSITE; PS00843; DDLA_DALA_LIGASE_1; 1.
 DR PROSITE; PS00843; DDLA_DALA_LIGASE_1; 1.
 KW Ligase; Complete proteome.
 SQ SEQUENCE 428 AA; 46045 MW; D501772E024C8F46 CRC64;

Query Match 91.7%; Score 33; DB 16; Length 428;
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 2.70175 Seconds
(without alignments)

817.479 Million cell updates/sec

Title: US-09-832-929-18_COPY_560_566

Perfect score: 36

Sequence: 1 KADDKET 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	396	4 Q81UK7	Q81UK7 homo sapien
2	36	100.0	417	4 Q86YG0	Q86YG0 homo sapien
3	36	100.0	598	3 O34674	O34674 schizosacch
4	33	91.7	402	10 Q80595	Q80595 arabidopsis
5	33	91.7	428	16 Q8G7C4	Q8G7C4 bifidobacte
6	33	91.7	673	5 Q95YMO	Q95YMO polyandroca
7	31	86.1	281	13 Q7TOR9	Q7TOR9 xenopus lae
8	31	86.1	341	16 Q9PP97	Q9PP97 campylobact
9	31	86.1	405	16 Q931U9	Q931U9 staphylococ
10	31	86.1	441	2 Q7X231	Q7X231 staphylococ
11	31	86.1	441	2 Q7X230	Q7X230 staphylococ
12	31	86.1	506	4 Q81YV1	Q81YV1 homo sapien
13	31	86.1	520	16 Q8CPR1	Q8CPR1 staphylococ
14	31	86.1	620	5 Q9GRW3	Q9GRW3 glottidia p
15	31	86.1	632	4 Q96NK7	Q96NK7 homo sapien
16	31	86.1	662	4 Q96M02	Q96M02 homo sapien

17	31	86.1	690	4 Q72353	Q72353 homo sapien
18	31	86.1	752	11 Q55035	Q55035 rattus norv
19	31	86.1	873	5 Q23717	Q23717 cryptospori
20	31	86.1	965	5 Q8T3I1	Q8T3I1 drosophila
21	31	86.1	970	10 Q22941	Q22941 arabidopsis
22	31	86.1	1214	16 Q87YMS	Q87YMS pseudomonas
23	31	86.1	1217	5 Q9VK44	Q9VK44 drosophila
24	31	86.1	1267	16 Q8A5K9	Q8A5K9 bacteroides
25	31	86.1	2361	5 Q8IEJ4	Q8IEJ4 plasmodium
26	30	83.3	161	16 Q88ZT1	Q88ZT1 lactobacill
27	30	83.3	527	17 Q9VBL7	Q9VBL7 aeropyrum p
28	30	83.3	1211	11 Q35233	Q35233 mus musculu
29	29	80.6	101	16 Q7V388	Q7V388 prochloroco
30	29	80.6	149	16 Q831T4	Q831T4 enterococcu
31	29	80.6	302	16 Q989T0	Q989T0 rhizobium l
32	29	80.6	364	16 Q9X277	Q9X277 thermotoga
33	29	80.6	368	17 Q8TN41	Q8TN41 methanosarc
34	29	80.6	377	5 Q21995	Q21995 caenorhabdi
35	29	80.6	397	16 Q7UHV8	Q7UHV8 rhodopirell
36	29	80.6	399	2 Q9ACN3	Q9ACN3 streptococc
37	29	80.6	431	16 Q97DR5	Q97DR5 clostridium
38	29	80.6	445	16 Q8XNA1	Q8XNA1 clostridium
39	29	80.6	447	17 Q8TSY4	Q8TSY4 methanosarc
40	29	80.6	454	16 Q9XAJ0	Q9XAJ0 streptomyce
41	29	80.6	649	16 Q8EX47	Q8EX47 mycoplasma
42	29	80.6	658	5 Q8T869	Q8T869 dictyosteli
43	29	80.6	684	5 Q9YXS2	Q9YXS2 drosophila
44	29	80.6	755	5 Q9N4K2	Q9N4K2 caenorhabdi
45	29	80.6	767	5 Q9U234	Q9U234 caenorhabdi

ALIGNMENTS

RESULT 1

Q81UK7 PRELIMINARY; PRT; 396 AA.
 ID Q81UK7
 AC Q81UK7;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Similar to serum albumin precursor.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_taxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC035969; AAH35969.1; .
 DR GO; GO:0005615; C:extracellular space; IEA.
 DR GO; GO:0005386; F:carrier activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR00264; Serum_albumin.
 DR Pfam; PF00273; transport_prot; 2.
 DR PRINTS; PRO0802; SERUMALBUMIN.
 DR SMART; SM00103; ALBUMIN; 2.
 DR PROSITE; PS00212; ALBUMIN; 2.
 SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;

Query Match 100.0%; Score 36; DB 4; Length 396;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
 |||||
 Db 371 KADDKET 377

RESULT 2

Q86YG0


```
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-604A-5
Query Match      100.0%; Score 36; DB 14; Length 585;
Best Local Similarity 100.0%; Pred.No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KADDKET 7
Db      560 KADDKET 566

RESULT 15
US-10-319-263-1
; Sequence 1, Application US/10319263
; Publication No. US20030180820A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISC007
; CURRENT APPLICATION NUMBER: US/10/319,263
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 60/115,992
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-319-263-1
Query Match      100.0%; Score 36; DB 14; Length 585;
Best Local Similarity 100.0%; Pred.No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KADDKET 7
Db      560 KADDKET 566

Search completed: April 19, 2004, 12:55:00
Job time : 2.94737 secs
```

Query Match 100.0%; Score 36; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
|||
Db 560 KADDKET 566

RESULT 10
US-10-424-999-11
; Sequence 11, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Arogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-11

Query Match 100.0%; Score 36; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
|||
Db 560 KADDKET 566

RESULT 11
US-10-425-000-31
; Sequence 31, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-31

Query Match 100.0%; Score 36; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
|||
Db 560 KADDKET 566

Db 560 KADDKET 566
|||
|||

RESULT 12
US-10-433-108-34
; Sequence 34, Application US/10433108
; Publication No. US20040053370A1
; GENERAL INFORMATION:
; APPLICANT: Eli Lilly and Company
; TITLE OF INVENTION: GLP-1 FUSION PROTEINS
; FILE REFERENCE: X-13991
; CURRENT APPLICATION NUMBER: US/10/433,108
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: US 60/251,954
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-433-108-34

Query Match 100.0%; Score 36; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
|||
Db 560 KADDKET 566

RESULT 13
US-10-153-064-5
; Sequence 5, Application US/10153064
; Publication No. US20020142814A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5

Query Match 100.0%; Score 36; DB 13; Length 585;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
|||
Db 560 KADDKET 566

RESULT 14
US-10-153-604A-5
; Sequence 5, Application US/10153604A
; Publication No. US20030143191A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,604A
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212

; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-117-18
Query Match 100.0%; Score 36; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KADDKET 7
DB 560 KADDKET 566
RESULT 6
US-09-932-322-445
; Sequence 445, Application US/09932322
; Publication No. US20030194743A1
; GENERAL INFORMATION:
; APPLICANT: Dyax Corp.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Ladner, Robert Charles
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BLYS)
; FILE REFERENCE: DYX-018.1 PCT; DYX-018.1 US
; CURRENT APPLICATION NUMBER: US/09/932,322
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-932-322-445
Query Match 100.0%; Score 36; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KADDKET 7
DB 560 KADDKET 566
RESULT 7
US-09-832-501-18
; Sequence 18, Application US/09832501
; Publication No. US20030199043A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J.
; APPLICANT: Sleep, Darrell
; APPLICANT: Turner, Andrew J.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF542
; CURRENT APPLICATION NUMBER: US/09/832,501
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-832-501-18
Query Match 100.0%; Score 36; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KADDKET 7
DB 560 KADDKET 566
RESULT 8
US-09-833-118-18
; Sequence 18, Application US/09833118
; Publication No. US20030219875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF544
; CURRENT APPLICATION NUMBER: US/09/833,118
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-118-18
Query Match 100.0%; Score 36; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KADDKET 7
DB 560 KADDKET 566
RESULT 9
US-09-833-245-18
; Sequence 18, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-245-18

```
;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-929-552-2

Query Match      100.0%; Score 36; DB 9; Length 585;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
Db 560 KADDKET 566

RESULT 2
US-09-932-613-445
; Sequence 445, Application US/09932613
; Publication No. US20030091565A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON
; FILE REFERENCE: DXX-025.1 PCT; DXX-025.1 US
; CURRENT APPLICATION NUMBER: US/09/932.613
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-932-613-445

Query Match      100.0%; Score 36; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
Db 560 KADDKET 566

RESULT 3
US-09-984-010-26
; Sequence 26, Application US/09984010
; Publication No. US20030104578A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David James
; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,010
; FILING DATE: 21-May-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/091,873
; FILING DATE: 25-JUN-1998
```

```
;
; APPLICATION NUMBER: PCT/GB96/03164
; FILING DATE: 19-DEC-1996
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-984-010-26

Query Match      100.0%; Score 36; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
Db 560 KADDKET 566

RESULT 4
US-09-833-041-18
; Sequence 18, Application US/09833041
; Publication No. US20030125247A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PFS45
; CURRENT APPLICATION NUMBER: US/09/833,041
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-041-18

Query Match      100.0%; Score 36; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
Db 560 KADDKET 566

RESULT 5
US-09-833-117-18
; Sequence 18, Application US/09833117
; Publication No. US20030171267A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PFS43
; CURRENT APPLICATION NUMBER: US/09/833,117
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 12:00:25 ; Search time 2.94737 Seconds

(without alignments)
654.724 Million cell updates/sec

Title: US-09-832-929-18_COPY_560_566

Perfect score: 36

Sequence: 1 KADDKET 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/PCT_NEW PUB. pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW PUB. pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06 NEW PUB. pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06 PUBCOMB. pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07 NEW PUB. pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS PUBCOMB. pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08 NEW PUB. pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08 PUBCOMB. pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A PUBCOMB. pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B PUBCOMB. pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C PUBCOMB. pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09 NEW PUB. pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A PUBCOMB. pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B PUBCOMB. pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C PUBCOMB. pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10 NEW PUB. pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60 NEW PUB. pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB. pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	36	100.0	585	9	US-09-929-552-2
2	36	100.0	585	10	US-09-932-613-445
3	36	100.0	585	10	US-09-984-010-26
4	36	100.0	585	10	US-09-833-041-18
5	36	100.0	585	10	US-09-833-117-18
6	36	100.0	585	10	US-09-932-322-445
7	36	100.0	585	10	US-09-832-501-18
8	36	100.0	585	11	US-09-833-118-18
9	36	100.0	585	11	US-09-833-245-18
10	36	100.0	585	12	US-10-424-999-11
11	36	100.0	585	12	US-10-425-000-31
12	36	100.0	585	12	US-10-433-108-34
13	36	100.0	585	13	US-10-153-064-5
14	36	100.0	585	14	US-10-153-604A-5
15	36	100.0	585	14	US-10-319-263-1

ALIGNMENTS

RESULT 1

US-09-929-552-2
; Sequence 2, Application US/09929552
; Patent No. US20020123080A1
; GENERAL INFORMATION:
; APPLICANT: Sonnschein, Carlos
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/929,552
; FILING DATE: 14-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/769,746
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBR1-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids

16	36	100.0	585	14	US-10-319-263-2	Sequence 2, Appli
17	36	100.0	585	14	US-10-414-469-1	Sequence 1, Appli
18	36	100.0	585	14	US-10-414-469-2	Sequence 2, Appli
19	36	100.0	585	14	US-10-413-831-1	Sequence 1, Appli
20	36	100.0	585	14	US-10-413-831-2	Sequence 2, Appli
21	36	100.0	585	15	US-10-413-832-1	Sequence 1, Appli
22	36	100.0	585	15	US-10-413-832-2	Sequence 2, Appli
23	36	100.0	585	15	US-10-414-386-1	Sequence 1, Appli
24	36	100.0	585	15	US-10-414-386-2	Sequence 2, Appli
25	36	100.0	585	15	US-10-233-675A-11	Sequence 11, Appl
26	36	100.0	585	15	US-10-462-262-26	Sequence 26, Appl
27	36	100.0	604	10	US-09-984-010-7	Sequence 7, Appli
28	36	100.0	609	10	US-09-919-039-370	Sequence 370, App
29	36	100.0	609	12	US-10-603-346-12	Sequence 12, Appl
30	36	100.0	609	13	US-10-153-084-7	Sequence 7, Appli
31	36	100.0	609	14	US-10-153-604A-7	Sequence 7, Appli
32	36	100.0	609	14	US-10-365-623-23	Sequence 23, Appl
33	36	100.0	610	9	US-09-984-186-2	Sequence 2, Appli
34	36	100.0	610	14	US-10-237-667-2	Sequence 2, Appli
35	36	100.0	610	14	US-10-237-708-2	Sequence 2, Appli
36	36	100.0	610	14	US-10-237-866-2	Sequence 2, Appli
37	36	100.0	610	14	US-10-237-871-2	Sequence 2, Appli
38	36	100.0	610	14	US-10-237-624-2	Sequence 2, Appli
39	36	100.0	616	12	US-10-433-108-13	Sequence 13, Appl
40	36	100.0	624	12	US-10-433-108-16	Sequence 16, Appl
41	36	100.0	631	12	US-10-433-108-14	Sequence 14, Appl
42	36	100.0	640	12	US-10-433-108-15	Sequence 15, Appl
43	36	100.0	640	12	US-10-433-108-17	Sequence 17, Appl
44	36	100.0	651	13	US-10-153-084-133	Sequence 133, App
45	36	100.0	651	14	US-10-153-604A-133	Sequence 133, App

Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
| | | | |
Db 560 KADDKET 566

RESULT 14
AAR26362
ID AAR26362 standard; protein; 585 AA.
XX AC AAR26362;
XX
DT 11-FEB-1993 (first entry)
XX
DE Synthetic HSA protein.
XX
KW Human serum albumin; transformants; recombinant.
XX
OS Synthetic.
XX
PN JP04211375-A.
XX
PD 03-AUG-1992.
XX
PF 05-FEB-1991; 91JP-00014600.
XX
PR 05-FEB-1990; 90JP-00025682.
XX
PA (AJIN) AJINOMOTO KK.
XX
DR WPI; 1992-304940/37.
DR N-PSDB; AAQ27813.
XX
XX Synthetic gene for prepn. of human serum albumin - comprises synthetic
PT DNA contg. gene coding the albumin using coding in Escherichia coli.
XX Fig 1; Page 13; 37pp; Japanese.
XX The protein sequence was deduced from the synthetic DNA sequence encoding
CC human serum albumin which was prepd. by ligating eight synthetic HSA gene
CC fragments
XX
SQ Sequence 585 AA;

Query Match 100.0%; Score 36; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
| | | | |
Db 560 KADDKET 566

RESULT 15
AAR20029
ID AAR20029 standard; protein; 585 AA.
XX AC AAR20029;
XX
DT 08-APR-1992 (first entry)
XX
DE Human serum albumin.
XX
KW HSA; yeast promoter; His4; Leu2.
XX
OS Homo sapiens.
XX
PN JF03262487-A.
XX
PD 22-NOV-1991.
XX

PF 12-MAR-1990; 90JP-00057885.
XX
PR 12-MAR-1990; 90JP-00057885.
XX (TOFU) TONEN CORP.
XX
DR WPI; 1992-012704/02.
DR N-PSDB; AAQ20201.
XX
PT Stable prepn. of human serum albumin - by culturing yeast in which
PT plasmid for recombinating DNA coding human serum albumin, etc. is inserted.
XX
PS Disclosure; Fig 4; 12pp; Japanese.
XX
CC HSA can be recombinantly produced in yeast. A HSA coding sequence is
CC incorporated into a vector which also contains a DNA region having the
CC same base sequence as that of the target region of the chromosome and a
CC promoter (esp. HIS4 or LEU2) to regulate the expression of the HSA coding
CC sequence in the yeast host
XX
SQ Sequence 585 AA;

Query Match 100.0%; Score 36; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
| | | | |
Db 560 KADDKET 566

Search completed: April 19, 2004, 11:51:22
Job time : 4.11727 secs

XX Human normal serum albumin A; pAT-pho-HSA-A; haemorrhagic shock;
KW hypoalbuminaemia.
XX
XX Homo sapiens.
XX OS
XX PN EP330451-A.
XX PD 30-AUG-1989.
XX PF 22-FEB-1989; 89EP-00301731.
XX PR 22-FEB-1988; 88JP-00037453.
XX PA (TOFU) TOA NENRYO KOGYO KK.
XX PI Suzuki M, Maki N, Yagi S;
XX WPI; 1989-250534/35.
XX CDNA encoding human normal serum albumin contained in plasmid - obtd. by
PT culturing host transformed with expression vector comprising CDNA coding
PT for albumin.
XX
XX Claim 1; Fig 3-1 to 3-5; 19pp; English.
XX
XX cDNA amino acid sequence of human serum albumin A (HSA-A) which is
CC identical to that encoded by chromosomal DNA. Previous polypeptides
CC produced from CDNA have one or more amino acids which differ from those
CC of HSA-A produced from the chromosomal DNA, and may exhibit antigenicity
CC when administered to humans. The HSA-A is used to treat haemorrhagic
CC shock and hypoalbuminaemia. See also AAN90600. (Updated on 25-MAR-2003 to
CC correct PI field.)
XX
XX Sequence 585 AA;
SQ
Query Match 100.0%; Score 36; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KADDKET 7
DB 560 KADDKET 566
RESULT 12
AAR05318
ID AAR05318 standard; protein; 585 AA.
AC
XX AAR05318;
XX 08-OCT-1990 (first entry)
XX Human serum albumin gene product.
XX Human serum albumin; HSA-A; yeast; ds.
XX Homo sapiens.
XX JP02117384-A.
XX 01-MAY-1990.
XX 26-OCT-1988; 88JP-00268302.
XX 26-OCT-1988; 88JP-00268302.
XX (TOFU) TOA NENRYO KOGYO KK.
XX WPI; 1990-176228/23.
XX N-PSDB; AAQ04719.
XX Human serum albumin prepn. by yeast host - by culturing transformed

PT plasmid yeast to produce serum, and removing it.
XX
XX Disclosure; Page 7; -pp; Japanese.
XX
XX Mature HSA-A may be produced using the sequence incorporated into a
CC plasmid vector with suitable controllers, and transferred to a yeast
CC expression system
XX
XX Sequence 585 AA;
SQ
Query Match 100.0%; Score 36; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KADDKET 7
DB 560 KADDKET 566
RESULT 13
AAR08457
ID AAR08457 standard; protein; 585 AA.
AC
XX AAR08457;
XX 25-MAR-2003 (revised)
XX 16-APR-1991 (first entry)
XX Human serum albumin.
XX HSA; folding; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT Region 1..303
FT /label= B
FT Region 123..585
FT /label= C
FT Region 123..303
FT /label= A
XX JP02227079-A.
XX PD 10-SEP-1990.
XX 25-AUG-1989; 89JP-00217540.
XX 06-OCT-1988; 88JP-00250926.
XX (TOFU) TONEN CORP.
XX WPI; 1990-317325/42.
XX N-PSDB; AAQ06099.
XX New human serum albumin fragments - used to bond to medicines and for
PT stable folding of protein(s).
XX Claim 1; Fig 8; 24pp; Japanese.
XX
XX Fragments A-C of HSA are expressed as fusion proteins with the signal
CC peptide of E. coli alkaline phosphatase. The fragments are selected for
CC their specific properties. The C-terminal truncated fragment, B, does not
CC bind long-chain fatty acids but does bind to various medicines at the
CC central region. The N-terminal truncated fragment, C, has good stability
CC in protein folding. The central segment, A, has characteristics of both B
CC and C. See also AAQ06096-Q06098. (Updated on 25-MAR-2003 to correct PD
CC field.) (Updated on 25-MAR-2003 to correct PF field.) (Updated on 25-MAR-
XX 2003 to correct PR field.)
XX Sequence 585 AA;
SQ
Query Match 100.0%; Score 36; DB 2; Length 585;

PT antibody specific to human albumin receptor.

PS Disclosure; Fig 2; 17pp; English.

XX The invention relates to testing human cancer cells, comprising obtaining cancer cells from the patient and contacting the cell *ex vivo* with an antibody to the receptor for human albumin. The method is useful for testing human cancer cells in particular breast and prostate cancer cells. The present sequence is mature human serum albumin, HSA. The anti-proliferative effect of HSA was assayed in an experiment included in the specification

XX Sequence 584 AA;

SQ

Query Match 100.0%; Score 36; DB 6; Length 584;
 Best Local Similarity 100.0%; Pred. No. 60;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KADDKET 7
 Db 559 KADDKET 565

RESULT 9

ID AAP93344 standard; protein; 585 AA.

XX AAP93344;

AC

XX 25-MAR-2003 (revised)

DT 23-JUN-1990 (first entry)

DE Sequence of mature human serum albumin (HSA) as encoded by artificial gene.

XX Mature human serum albumin; artificial gene; oligonucleotide block; hypobolaemia; shock; hypoalbuminaemia.

KW Homo sapiens.

OS

XX EP308381-A.

PN

XX 22-MAR-1989.

PD

XX 13-SEP-1988; 88EP-00850299.

PF

XX 14-SEP-1987; 87SE-00003539.

PR

XX (SKAN-) SKANDIGEN AB.

PA (MAGY) MTA SZEGEDI BIOLOG KOEPPONTI.

PA (VEPE-) VEPEX CONTRACTOR LTD.

XX

PI Aberg B, Simoncsits A, Kalan M, Cseprian I, Bajszar G;

XX

DR WPI; 1989-087749/12.

DR N-PSDB; AAN90997.

XX

PT Artificial gene coding for authentic human serum albumin - constructed on the basis of codons most frequently used by chosen non-human host.

PS Disclosure; pp. 11-16; 121pp; English.

XX The synthetic gene was constructed by designing a nucleotide sequence in which the codons which are most frequently used by the chosen non-human host were selected. In this case, it is yeast cells (LL20; Leu2-3; 112, His 3-11, 15). The synthetic HSA gene was assembled from 24 oligonucleotide blocks. HSA is used in therapy for the treatment of hypovolaemia, shock and hypoalbuminaemia. It is also used as an additive in perfusion liq. for extracorporeal circulation and as an experimental antigen. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 585 AA;

SQ

Query Match 100.0%; Score 36; DB 1; Length 585;
 Best Local Similarity 100.0%; Pred. No. 60;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KADDKET 7
 Db 560 KADDKET 566

RESULT 10

ID AAP90388 standard; protein; 585 AA.

XX AAP90388;

AC

XX 24-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 01-NOV-1989 (first entry)

XX Mature human serum albumin polypeptide.

DE Human serum albumin; mature protein; new polypeptides; plasma expanders.

KW Homo sapiens; (Human).

OS

XX EP322094-A.

PN

XX 28-JUN-1989.

PD

XX 25-OCT-1988; 88EP-00310000.

PF

XX 30-OCT-1987; 87GB-00025529.

PR

XX (DELZ) DELTA BIOTECHNOLOGY LTD.

PA

XX Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;

PI WPI; 1989-186464/26.

DR N-PSDB; AAN90128.

DR

XX New N-terminal fragments of human serum albumin - esp. useful as blood plasma expanders.

PT

XX Disclosure; Fig 2; 20pp; English.

PS

XX Mature protein of human serum albumin (see corresp. AAN90128). Used to make new N-terminal fragments which are used as plasma expanders, or as substitutes for HSA or BSA, in tissue culture media. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS field)

XX

SQ Sequence 585 AA;

Query Match 100.0%; Score 36; DB 1; Length 585;
 Best Local Similarity 100.0%; Pred. No. 60;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KADDKET 7
 Db 560 KADDKET 566

RESULT 11

ID AAP91422 standard; protein; 585 AA.

XX AAP91422;

AC

XX 25-MAR-2003 (revised)

DT 20-DEC-1989 (first entry)

XX Human normal serum albumin A.

DE

CC the process for producing human serum albumin in the yeast host cell,
CC especially in secretory mode
XX
SQ Sequence 228 AA;

Query Match 100.0%; Score 36; DB 3; Length 228;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
| | | | |
DB 203 KADDKET 209

RESULT 4

AAU33087
ID AAU33087 standard; protein; 243 AA.
XX
AC AAU33087;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #3578.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

OS Homo sapiens.

PN WO200179449-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001WO-US008656.

PR 18-APR-2000; 2000US-00552929.

PR 26-JAN-2001; 2001US-00770160.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-611725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy.

XX Claim 20; Page 706; 765pp; English.

CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU3304 represent the amino acid sequences of novel human
CC secreted proteins of the invention

XX Sequence 243 AA;

Query Match 100.0%; Score 36; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
| | | | |
DB 220 KADDKET 226

RESULT 5

AAU29876
ID AAU29876 standard; protein; 401 AA.

XX AAU29876;

XX 18-DEC-2001 (first entry)

XX Novel human secreted protein #367.

XX Human; vaccination; gene therapy; nutritional supplement;
XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;
XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.

OS Homo sapiens.

XX WO200179449-A2.

XX 25-OCT-2001.

XX 16-APR-2001; 2001WO-US008656.

XX 18-APR-2000; 2000US-00552929.

XX 26-JAN-2001; 2001US-00770160.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-611725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy.

XX Claim 20; Page 206; 765pp; English.

CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU3304 represent the amino acid sequences of novel human
CC secreted proteins of the invention

XX Sequence 401 AA;

Query Match 100.0%; Score 36; DB 4; Length 401;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KADDKET 7
| | | | |
DB 376 KADDKET 382

RESULT 6

AAU14179

PS Claim 10; SEQ ID NO 198; 153pp; English.

XX Sequences AAU27676-AAU28019 represent full-length polypeptides and contig

CC polypeptides of the invention. The proteins and their associated DNA

CC sequences are useful for the treatment, diagnosis and prevention of

CC various types of disorder in a mammalian subject such as a human, dog,

CC monkey, mouse, hamster or rat. The disorders include cancers such as

CC leukaemia, lymphoma and neuroblastoma, autoimmune disorders such as

CC multiple sclerosis, connective tissue disease, rheumatoid arthritis,

CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system

CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's

CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and

CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's

CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory

CC bowel disease. The sequences exhibit activity relating to angiogenesis,

CC cell proliferation, cell differentiation, stem cell growth factor,

CC activin or inhibin. Therefore, they can be used to manipulate stem cells

CC in culture to give rise to neuroepithelial cells that can be used to

CC augment or replace cells damaged by illness, accidental damage or genetic

CC disorders. The sequences may also be used for regeneration of bone,

CC cartilage, tendons and ligaments and in tissue repair and burn healing.

CC Note: Some sequences for this patent did not form part of the printed

CC specification, but were obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 140 AA;

XX

Query Match 100.0%; Score 36; DB 4; Length 140;

Best Local Similarity 100.0%; Pred. NO. 14;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KADDKET 7

DB 115 KADDKET 121

RESULT 2

AAU33289

ID AAU33289 standard; protein; 212 AA.

XX AAU33289;

XX

DT 18-DEC-2001 (first entry)

XX

DE Novel human secreted protein #3780.

XX

KW Human; vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX

OS Homo sapiens.

XX

PN WO200179449-A2.

XX

PD 25-OCT-2001.

XX

PF 15-APR-2001; 2001WO-US008656.

XX

PR 18-APR-2000; 2000US-00552929.

PR

XX 26-JAN-2001; 2001US-00770160.

XX

PA (HYSE-) HYSEQ INC.

XX

XX Tang YT, Liu C, Drmanac RT;

XX

XX WPI; 2001-611725/70.

XX

DR Nucleic acids encoding a range of human polypeptides, useful in genetic

PT vaccination, testing and therapy.

PT

XX Claim 20; Page 755; 765pp; English.

PS

XX The invention relates to novel human secreted polypeptides. The

CC

CC polypeptides and antibodies to the polypeptides are useful for

CC determining the presence of or predisposition to a disease associated

CC with altered levels of polypeptide. The polypeptides are also useful for

CC identifying agents (agonists and antagonists) that bind to them. Cells

CC expressing the proteins are useful for identifying a therapeutic agent

CC for use in treatment of a pathology related to aberrant expression or

CC physiological interactions of the polypeptide. Vectors comprising the

CC nucleic acids encoding the polypeptides and cells genetically engineered

CC to express them are also useful for producing the proteins. The proteins

CC are useful in genetic vaccination, testing and therapy, and can be used

CC as nutritional supplements. They may be used to increase stem cell

CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon

CC and/or nerve tissue growth or regeneration; immune suppression and/or

CC stimulation, as anti-inflammatory agents; and in treatment of leukaemias.

CC AAU29510-AAU3304 represent the amino acid sequences of novel human

CC secreted proteins of the invention

XX

XX Sequence 212 AA;

XX

Query Match 100.0%; Score 36; DB 4; Length 212;

Best Local Similarity 100.0%; Pred. NO. 22;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KADDKET 7

DB 187 KADDKET 193

RESULT 3

AAU83949

ID AAU83949 standard; protein; 228 AA.

XX AAU83949;

XX

DT 28-JUL-2000 (first entry)

XX

DE Yeast codon-biased recombinant HSA protein fragment HSA-III.

XX

KW Recombinant; human serum albumin; HSA; yeast codon bias; host cell;

KW overlapping oligonucleotide; expression vector.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN CN1239103-A.

XX

PD 22-DEC-1999.

XX

PF 17-JUN-1998; 98CN-00102506.

XX

PR 17-JUN-1998; 98CN-00102506.

XX

XX (HAIJ-) HAIJI BIOENGINEERING CO LTD.

XX

XX Li S, Lu D;

XX

PI WPI; 2000-351198/31.

DR

DR N-PSDB; AAA10094.

XX

PT Process for preparing recombinant human serum albumin comprising yeast

PT biased sex codons - uses a recombinant DNA technique.

XX

XX Example 1; Fig 7; 44pp; Chinese.

XX

XX The method relates to a method of recombinantly producing human serum

CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise

CC a yeast codon bias. The complete HSA gene (AAA10091) was generated as

CC three synthetic fragments (AAA10092-AA10094) joined by recombinant DNA

CC technology. Each HSA fragment was synthesised from overlapping

CC oligonucleotide fragments that were extended. This sequence represents

CC the sequence of the HSA fragment HSA-III encoded by the human gene with a

CC yeast codon bias. The invention also covers a recombinant expression

CC vector, yeast host cells carrying the recombinant expression vector and

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:24:29 ; Search time 4.11727 Seconds
(without alignments)
480.375 Million cell updates/sec

Title: US-09-832-929-18_COPY_560_566

Perfect score: 36

Sequence: 1 KADDKET 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	140	AAU27701	Aau27701 Human ful
2	36	100.0	212	AAU33289	Aau33289 Human hum
3	36	100.0	228	AAU83949	Aay83949 yeast cod
4	36	100.0	243	AAU33087	Aau33087 Novel hum
5	36	100.0	401	AAU29876	Aau29876 Novel hum
6	36	100.0	463	AAU14179	Aar14179 Human ser
7	36	100.0	550	AAU29877	Aau29877 Novel hum
8	36	100.0	584	ABG72381	Abg72381 Mature hu
9	36	100.0	585	AAU93344	Aap93344 Sequence
10	36	100.0	585	AAU90388	Aap90388 Mature hu
11	36	100.0	585	AAU91422	Aap91422 Human nor
12	36	100.0	585	AAU05318	Aar05318 Human ser
13	36	100.0	585	AAU08457	Aar08457 Human ser
14	36	100.0	585	AAU26362	Aar26362 Synthetic
15	36	100.0	585	AAU20029	Aar20029 Human ser
16	36	100.0	585	AAU80301	Aar80301 Human ser
17	36	100.0	585	AAU02011	Aao02011 HSA prote
18	36	100.0	585	AAU59841	Aaw59841 Mature pr
19	36	100.0	585	AAU84873	Aay84873 Amino aci
20	36	100.0	585	AAU83946	Aay83946 Yeast cod
21	36	100.0	585	AAU52567	Aam52567 Mature hu
22	36	100.0	585	AAU12417	Aae12417 Human alb
23	36	100.0	585	AAU12403	Aae12403 Human alb
24	36	100.0	585	AAU13129	Aae13129 Human alb
25	36	100.0	585	AAU13135	Aae13135 Human alb

ALIGNMENTS

RESULT 1

AAU27701
ID AAU27701 standard; protein; 140 AA.

AC AAU27701;

DT 18-DEC-2001 (first entry)

DE Human full-length polypeptide sequence #26.

XX Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
KW nervous system disorder; inflammatory disorder; cell differentiation;
KW angiogenesis; stem cell growth factor; actin; inhibitor; cartilage; burn;
KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
KW cytoskeletal; antirheumatic; antiarthritic; vulvular; antiparkinsonian;
KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
KW neuroprotective; osteopathic; antidiabetic; antiallergic;
KW immunostimulant; analgesic; gene therapy.

XX Homo sapiens.

OS WO200164834-A2.

PD 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004926.

PR 28-FEB-2000; 2000US-00515126.

PR 18-MAY-2000; 2000US-00577409.

PR 17-JUN-2000; 2000US-00597707.

PR 14-JUL-2000; 2000US-00616807.

PR 19-SEP-2000; 2000US-00664641.

PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QX, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;

PI Drmanac R;

XX WPI; 2001-589862/66.

DR N-PSDB; RAS44601.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries

PT prepared from various human tissues, for diagnosis, treatment of cancer,
PT neurological, inflammatory disorders and for use in arrays for detection.

XX

A;Cross-references: GB:AE004618; GB:AE004091; NID:99947912; PIDN:AG05311.1; GSPDB:GN001
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA1923

Query Match 75.6%; Score 34; DB 2; Length 1281;
Best Local Similarity 75.0%; Pred. No. 75;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 TESLVNRRP 9
|:|:|:|:|
DB 1269 QSVNRRP 1276

RESULT 12
J80106
hypochemical 11K protein (col-ND4L intergenic region) - Podospora anserina mitochondrion
C;Species: mitochondrion Podospora anserina
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 07-Dec-1999
C;Accession: J80106
R;Vierny-Jamet, C.
Gene 74, 387-398, 1988
A;Title: Senescence in Podospora anserina: a possible role for nucleic acid interacting
structures.
A;Reference number: P80026; MUID:89232730; PMID:3246349
A;Accession: J80106
A;Molecule type: DNA
A;Residues: 1-97 <VIE>
A;Note: This reading frame extends between two stop codons and does not begin with a sta
C;Comment: this protein is encoded by senescence-specific DNA (sen-DNA), which is result
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC3
C;Keywords: mitochondrion

Query Match 73.3%; Score 33; DB 2; Length 97;
Best Local Similarity 75.0%; Pred. No. 6.9;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TESLVNRR 8
|:|:|:|:|
DB 7 TEELINRR 14

RESULT 13
A42391
Ca2+-transporting ATPase (EC 3.6.3.8) PMCA4b - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 19-Apr-2002
C;Accession: A42391
R;Brandt, P.; Neve, R.L.; Kamesheidt, A.; Rhoads, R.E.; Vanaman, T.C.
J. Biol. Chem. 267, 4376-4385, 1992
A;Title: Analysis of the tissue-specific distribution of mRNAs encoding the plasma membr
c levels.

A;Reference number: A42391; MUID:92165787; PMID:1531651
A;Accession: A42391
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-473 <BRA>
A;Cross-references: GB:M83364; NID:9163549; PIDN:AAA30713.1; PID:G163550
A;Note: sequence extracted from NCBI backbone (NCBIN:85263, NCBI:P:85264)
C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C;Keywords: ATP; hydrolase
F;1-153/Domain: ATPase nucleotide-binding domain homology (fragment) <ATN>

Query Match 73.3%; Score 33; DB 2; Length 473;
Best Local Similarity 66.7%; Pred. No. 41;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
|:|:|:|:|
DB 198 TDSLLKRRP 206

RESULT 14

S54356
Ca2+-transporting ATPase (EC 3.6.3.8), plasma membrane isoform 4a - rat
N;Alternate names: calcium pump form PMCA4a
C;Species: Rattus norvegicus (Norway rat)
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 19-Apr-2002
C;Accession: S54356; F44525
R;Keeton, T.P.; Shull, G.E.
Biochem. J. 306, 779-785, 1995
A;Title: Primary structure of rat plasma membrane Ca(2+)-ATPase isoform 4 and analysis
A;Reference number: S54356; MUID:95217154; PMID:7702574
A;Accession: S54356
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1169 <KEE>
A;Cross-references: EMBL:U15408; NID:G606965; PIDN:AAA81008.1; PID:G606966
R;Keeton, T.P.; Burk, S.E.; Shull, G.E.
J. Biol. Chem. 268, 2740-2748, 1993
A;Title: Alternative splicing of exons encoding the calmodulin-binding domains and C t
A;Reference number: A45213; MUID:93155089; PMID:8428948
A;Accession: F44525
A;Molecule type: DNA; mRNA
A;Residues: 1081-1169 <KE2>
A;Cross-references: GB:L05569; GB:L04742; NID:G206268; PIDN:AAA50820.1; PID:G206270
A;Note: sequence extracted from NCBI backbone (NCBI:P:124233)
C;Genetics:
A;Gene: PMCA4
C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C;Keywords: alternative splicing; ATP; calmodulin binding; hydrolase; membrane protein
F;666-851/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 73.3%; Score 33; DB 2; Length 1169;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
|:|:|:|:|
DB 896 TDSLLRRP 904

RESULT 15

S54357
plasma membrane Ca2+-ATPase isoform 4 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 22-Jun-1999
C;Accession: S54357
R;Keeton, T.P.; Shull, G.E.
Biochem. J. 306, 779-785, 1995
A;Title: Primary structure of rat plasma membrane Ca(2+)-ATPase isoform 4 and analysis
A;Reference number: S54356; MUID:95217154; PMID:7702574
A;Accession: S54357
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1203 <KEE>
A;Cross-references: EMBL:U15408; NID:G606965; PIDN:AAA81006.1; PID:G1054879
C;Genetics:
A;Introns: 1103/3
C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C;Keywords: alternative splicing; ATP
F;666-851/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 73.3%; Score 33; DB 2; Length 1203;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
|:|:|:|:|
DB 896 TDSLLRRP 904

Search completed: April 19, 2004, 12:02:33
Job time : 2.09695 secs

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 403-472 <BR2>
A>Note: sequence extracted from NCBI backbone (NCBIN:85265, NCIP:85266)
A:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C:Keywords: ATP; hydrolase
F:1-153/Domain: ATPase nucleotide-binding domain homology (fragment) <ATN>

Query Match 80.0%; Score 36; DB 2; Length 472;
Best Local Similarity 77.8%; Pred. No. 9;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
|||||
Db 198 TESLLKRP 206

RESULT 8
A35547
Ca2+-transporting ATPase (EC 3.6.3.8) 3, plasma membrane - human
N:Alternate names: Ca2+-transporting ATPase hPMCA3; Ca2+-transporting ATPase, erythrocyte
C:Species: Homo sapiens (man)
C:Date: 21-Sep-1990 #sequence_revision 21-Sep-1990 #text_change 19-Apr-2002
C:Accession: A35547; A29321; A29930; B31332
R:Strehler, E.E.; James, P.; Fischer, R.; Heim, R.; Vorherr, T.; Filoteo, A.G.; Penniston, J. Biol. Chem. 265, 2835-2842, 1990
A:Title: Peptide sequence analysis and molecular cloning reveal two calcium pump isoforms
A:Reference number: A35547; MUID:90153913; PMID:2137451
A:Accession: A35547
A:Molecule type: mRNA
A:Residues: 1-1205 <SR>
A:Cross-references: GB:M55874; NID:9179162; PIDN:AAA50819.1; PID:G179163
R:Filoteo, A.G.; Gorski, J.P.; Penniston, J.T.
J. Biol. Chem. 262, 6526-6530, 1987
A:Title: The ATP-binding site of the erythrocyte membrane Ca(2)-pump. Amino acid sequence
A:Reference number: A29321; MUID:87194687; PMID:2952652
A:Accession: A29321
A:Molecule type: protein
A:Residues: 588-599 <FLL>
R:James, P.; Maeda, M.; Fischer, R.; Verma, A.K.; Krebs, J.; Penniston, J.T.; Carafoli, J. Biol. Chem. 263, 2905-2910, 1988
A:Title: Identification and primary structure of a calmodulin binding domain of the Ca(2)-pump
A:Reference number: A29930; MUID:88139343; PMID:2963820
A:Accession: A29930
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1085-1107, 'N', 1109-1110, 'S', 1112-1116, 'F' <JAM>
R:Brandt, P.; Zurini, M.; Neve, R.L.; Rhoads, R.E.; Vanaman, T.C.
Proc. Natl. Acad. Sci. U.S.A. 85, 2914-2918, 1988
A:Title: A C-terminal, calmodulin-like regulatory domain from the plasma membrane Ca(2+)-pump
A:Reference number: A94710; MUID:88203611; PMID:2966397
A:Accession: B31332
A:Molecule type: protein
A:Residues: 1141, 'Q', 1143-1150, 'Q', 1152-1154 <BRA>
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C:Keywords: ATP; hydrolase; membrane protein; phosphoprotein
F:666-851/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 80.0%; Score 36; DB 2; Length 1205;
Best Local Similarity 77.8%; Pred. No. 26;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
|||||
Db 896 TESLLKRP 904

RESULT 9
T33877
hypothetical protein R05C11.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Aug-2000
C:Accession: T33877

R:Lamar, B.; Kramer, J.; Gibson, A.
submitted to the EMBL Data Library, February 1999
A:Description: The sequence of C. elegans cosmid R05C11.
A:Reference number: Z21429
A:Accession: T33877
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1110 <LAM>
A:Cross-references: EMBL:AF125446; PIDN:AAD12806.1; GSPDB:GN00022; CESP:R05C11.3
A:Experimental source: strain Bristol N3; clone R05C11
C:Genetics: CESP:R05C11.3
A:Map position: 4
A:Introns: 28/3; 115/1; 155/1; 358/3; 443/1; 776/1; 963/3; 1018/3; 1087/3
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
F:639-824/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 77.8%; Score 35; DB 2; Length 1110;
Best Local Similarity 66.7%; Pred. No. 39;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
|||||
Db 869 TEDLLNRP 877

RESULT 10
GB3162
respiratory nitrate reductase alpha chain PA3875 [imported] - Pseudomonas aeruginosa
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: GB3162
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Liradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Liradman, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: GB3162
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1261 <STO>
A:Cross-references: GB:AE004804; GB:AE004091; NID:G9950044; PIDN:AAG07262.1; GSPDB:GN004091
A:Experimental source: strain PA01
C:Genetics: CESP:PA3875
C:Superfamily: nitrate reductase alpha chain

Query Match 77.8%; Score 35; DB 2; Length 1261;
Best Local Similarity 66.7%; Pred. No. 45;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
|||||
Db 1077 TEKLLNRP 1085

RESULT 11
GB3405
hypothetical protein PA1923 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: GB3405
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Liradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Liradman, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: GB3405
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1281 <STO>

A:Title: Mass spectrometric identification of modifications to human serum albumin treated with hydrogen peroxide
A:Reference number: S36882; MUID:93384321; PMID:8373198
A:Accession: S36882
A:Molecule type: protein
A:Residues: 45-67;141-160;311-337;469-490;570-581 <FIN>
R:Kausler, E.; Spittler, G.
Biol. Chem. Hoppe-Seyler 372, 849-855, 1991
A:Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmolekule
A:Reference number: S17599; MUID:92126241; PMID:1772598
A:Accession: S17599
A:Molecule type: protein
A:Residues: 25-54;154-357;431-447 <KAU>
A:Note: 49-Leu was also found
R:Caraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A:Title: Structures of histamine-releasing peptides formed by the action of acid proteases
A:Reference number: A45800; MUID:89341406; PMID:2474609
A:Accession: A45800
A:Molecule type: protein
A:Residues: 166-173 <CAR>
R:Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; Wang, J.
Biochem. Biophys. Res. Commun. 136, 993-998, 1986
A:Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-treated porcine pancreas
A:Reference number: A03239; MUID:86242180; PMID:3087352
A:Accession: A03239
A:Molecule type: protein
A:Residues: 166-173, 'L' <MOG>
R:Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, S.
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990
A:Title: Mutations in genetic variants of human serum albumin found in Italy
A:Reference number: A38255; MUID:91062352; PMID:2247440
A:Accession: A38255
A:Molecule type: protein
A:Residues: 76-111 <GAL1>
A:Accession: B38255
A:Molecule type: protein
A:Residues: 82-105, 'K', 107-110 <GAL2>
A:Note: this variant is designated albumin Vilbo Valentia
A:Accession: A38255
A:Molecule type: protein
A:Residues: 76-83, 'K', 85-106 <GAL3>
A:Note: this variant is designated albumin Torino
R:Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.
Eur. J. Biochem. 214, 437-444, 1993
A:Title: The structural characterization and bilirubin-binding properties of albumin Her
A:Reference number: S33298; MUID:93292504; PMID:8513793
A:Accession: S33298
A:Molecule type: protein
A:Residues: 255-263, 'E', 265-281 <MIN1>
A:Note: this variant is designated albumin Herborn
R:Minchiotti, L.; Galliano, M.; Stoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta, F.
Biochim. Biophys. Acta 1119, 232-238, 1992
A:Title: Two albumins with identical electrophoretic mobility are produced by different genes
A:Reference number: S21078; MUID:92190239; PMID:1347703
A:Accession: S21078
A:Molecule type: protein
A:Residues: 354-356, 'K', 358-378 <MIN2>
A:Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported, R:He, X.M.; Carter, D.C.
Nature 358, 209-215, 1992
A:Title: Atomic structure and chemistry of human serum albumin
A:Reference number: A46756; MUID:92334427; PMID:1630489
A:Accession: A46756
A:Contents: annotation; X-ray crystallography, 2.8 angstroms
R:Brown, J.R.; Shokley, P.; Behrens, F.Q.
In The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40, 1977
A:Reference number: A94442
A:Contents: annotation; three-dimensional structure and disulfide bonds
R:Saber, M.A.; Stockbauer, P.; Moravsek, L.; Meloun, B.
Collect. Czech. Chem. Commun. 42, 564-579, 1977
A:Title: Disulfide bonds in human serum albumin
A:Reference number: A90930
A:Contents: annotation; disulfide bonds
R:Jacobsen, C.

Biochem. J. 171, 453-459, 1978
A:Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding of bilirubin
A:Reference number: A90299; MUID:78186630; PMID:656055
A:Contents: annotation; bilirubin-binding site
R:Peters, T.; Reed, R.G.
in: Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjolholm, I., eds., 11-20
A:Title: Serum albumin: conformation and active sites
A:Reference number: A94408
A:Contents: annotation; binding sites
R:Harper, M.E.; Dugaiczky, A.
Am. J. Hum. Genet. 35, 565-572, 1983
A:Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein gene
A:Reference number: A90028; MUID:83279982; PMID:6192711
A:Contents: annotation; gene position
R:Walker, J.E.
FEBS Lett. 66, 173-175, 1976
A:Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid
A:Reference number: A46755; MUID:76257808; PMID:955075
A:Contents: annotation
R:Bohney, J.P.; Fonda, M.L.; Feldhoff, R.C.
FEBS Lett. 298, 266-268, 1992
A:Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phosphate
A:Reference number: A56294; MUID:92183881; PMID:1544460
A:Contents: annotation
A:Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in
ataze activity
C:Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized
in liver, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak)
C:Comment: A large number of variants of human serum albumin have been described.
C:Genetics:
A:Gene: GDB:ALB
A:Cross-references: GDB:118990; OMIM:103600
A:Map position: 4q11-q13
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyrid:
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status experimental <PRO>
F:25-603/Product: serum albumin #status experimental <MPT>
F:166-174/Product: kinetensin #status experimental <SAI>
F:221-394/Domain: serum albumin repeat homology <KIP>
F:413-592/Domain: serum albumin repeat homology <SA2>
F:727/Binding site: copper (His) #status predicted
F:77-86;99-115;114-125;148-193;192-201;224-270;269-277;289-303;302-313;340-385;384-393,
F:214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

Query Match 100.0%; Score 45; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TESLVNRRP 9
DB 502 TESLVNRRP 510

RESULT 7
C42391
C42+transporting ATPase (EC 3.6.3.8) PMCA4b - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 19-Apr-2002
C:Accession: C42391; B42391
R:Brandt, P.; Neve, R.L.; Kammerheid, A.; Rhoads, R.E.; Vanaman, T.C.
J. Biol. Chem. 267, 4376-4385, 1992
A:Title: Analysis of the tissue-specific distribution of mRNAs encoding the plasma mem
c levels.
A:Reference number: A42391; MUID:92165787; PMID:1531651
A:Accession: C42391
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-472 <BRA>
A:Note: sequence extracted from NCBI backbone (NCBI:P:85267)
A:Accession: B42391


```

A:Molecule type: mRNA
A:Residues: 1-607 <BR>
A:Cross-references: EMBL:X17055; NID:G1386; PIDN:CAA34903.1; PID:G1387
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the membra
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-607/Product: serum albumin #status predicted <MAT>
F:29-201/Domain: serum albumin repeat homology <SA1>
F:220-393/Domain: serum albumin repeat homology <SA2>
F:412-591/Domain: serum albumin repeat homology <SA3>
F:27/Binding site: copper (His) #status predicted
F:77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4
F:263/Binding site: bilirubin (Lys) #status predicted

Query Match 100.0%; Score 45; DB 1; Length 607;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TESLVNRRP 9
|||||
Db 501 TESLVNRRP 509

RESULT 5
SS7632
serum albumin precursor - cat
C:Species: Felis silvestris catus (domestic cat)
C:Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Aug-1999
C:Accession: J04660; S57632
R:Hitler, C.; Grigoloni, F.; Hentges, P.
Gen 169, 295-296, 1996
A:Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.
A:Reference number: J04660; MUID:96194824; PMID:8647469
A:Accession: J04660
A:Molecule type: mRNA
A:Residues: 1-608 <HI>
A:Cross-references: EMBL:X84842; NID:G886484; PIDN:CAA59279.1; PID:G886485
A:Experimental source: liver
C:Comment: This protein is the major protein component in plasma. It functions as a mult
ein has 35 conserved cysteine residues.
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: liver; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRP>
F:25-608/Product: serum albumin #status predicted <MAT>
F:29-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>

Query Match 100.0%; Score 45; DB 2; Length 608;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TESLVNRRP 9
|||||
Db 502 TESLVNRRP 510

RESULT 6
ASBHS
serum albumin precursor [validated] - human
N:Alternate names: preproalbumin
N:Contains: kinetensin
C:Species: Homo sapiens (man)
C:Date: 29-Jul-1981 #sequence_revision 31-Jan-1997 #text_change 17-Mar-2000
C:Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S36
R:Law, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebur
Nucleic Acids Res. 9, 6103-6114, 1981
A:Title: The sequence of human serum albumin cDNA and its expression in Escherichia coli
A:Reference number: A93743; MUID:82081882; PMID:6171778

```

```

A:Accession: A93743
A:Molecule type: mRNA
A:Residues: 1-419, 'K', 421-609 <LAW>
A:Cross-references: EMBL:V00495; GB:J00078; GB:L00132; GB:L00133; NID:G28591; PIDN:CAA:
R:Dugaczky, A.; Law, S.W.; Dennison, O.E.
Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982
A:Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.
A:Reference number: A93936; MUID:82105994; PMID:6275391
A:Accession: A93936
A:Molecule type: mRNA
A:Residues: 1-120, 'G', 122-609 <DUG>
A:Cross-references: EMBL:V00494; NID:G28589; PIDN:CAA23753.1; PID:G28590
R:Urano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.
J. Biol. Chem. 261, 3244-3251, 1986
A:Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and
A:Reference number: I39427; MUID:86140099; PMID:2419329
A:Accession: I39427
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-26 <URA>
A:Cross-references: GB:M13075; NID:G178330; PIDN:AAA51688.1; PID:G553173
R:Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994
A:Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family
A:Reference number: I59286; MUID:94181575; PMID:8134387
A:Accession: I59286
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 282-290, 'KSRFDLQ' <WAT>
A:Cross-references: GB:S69192; NID:G546032; PIDN:AAB30282.1; PID:G546033
A:Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia
R:Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam,
Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994
A:Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl
A:Reference number: I59313; MUID:94294404; PMID:8022807
A:Accession: I59313
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 589-590, 'ALPRRVKNLLQVLP' <MAD>
A:Cross-references: GB:S70799; NID:G547231; PIDN:AAB31177.1; PID:G547232
A:Note: this frame-shift variant is designated albumin Bazzano; four additional variant
submitted to the EMBL Data Library, March 1995
R:Menaya, J.; Parrilla, R.; Ayuso, M.S.
A:Reference number: G08292
A:Accession: G01747
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-120, 'G', 122-455 <MEN>
A:Cross-references: EMBL:U22961; NID:G763428; PIDN:AAA64922.1; PID:G763431
R:Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.
Biochem. J. 308, 321-325, 1995
A:Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Ke
A:Reference number: S55314; MUID:95275251; PMID:7755581
A:Accession: S55314
A:Molecule type: protein
A:Residues: 19-27 <LED>
R:Meloun, B.; Moravek, L.; Kostka, V.
FEBS Lett. 58, 134-137, 1975
A:Title: Complete amino acid sequence of human serum albumin.
A:Reference number: A91420; MUID:76187907; PMID:1225573
A:Accession: A91420
A:Molecule type: protein
A:Residues: 25-117, 'EQ', 120-154, 'Q', 156-193, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395:
R:Roehr, U.; Spittler, G.; Triptier, D.
Justus Liebig's Ann. Chem. 9, 881-884, 1988
A:Title: Isolation and structure elucidation of middle-molecular weight peptides from
A:Reference number: S06422
A:Note: this paper is in German, with an English abstract
A:Accession: S06422
A:Molecule type: protein
A:Residues: 25-48 <ROS>
R:Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.
Arch. Biochem. Biophys. 305, 595-599, 1993

```

A;Molecule type: protein
A;Residues: 23-51, 'X', 53-54, 'XXXGV', 146, 'E', 148, 'E', 150-151, 'XVN', 155 <LIM>
A;Experimental source: dental enamel
A;Note: albumin and other serum proteins are also found in bone
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membra
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; Plasma
F;1-16/Domain: signal sequence (fragment) #status predicted <SIG>
F;17-22/Domain: signal sequence #status predicted <PRO>
F;23-605/Product: serum albumin #status predicted <MAT>
F;27-199/Domain: serum albumin repeat homology <SA1>
F;218-391/Domain: serum albumin repeat homology <SA2>
F;410-599/Domain: serum albumin repeat homology <SA3>
F;75-84, 97-113, 112-123, 145-190, 189-198, 221-267, 266-274, 286-300, 299-310, 337-382, 381-390, 4
F;261/Binding site: bilirubin (Lys) #status predicted

Query Match 100.0%; Score 45; DB 1; Length 605;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
|||||
Db 499 TESLVNRRP 507

RESULT 3
ABBO5
N;Alternate names: 67K protein; preproalbumin
C;Species: Bos primigenius taurus (cattle)
C;Date: 24-Apr-1984 #sequence revision 30-Sep-1993 #text change 18-Aug-2000
A;Accession: A38885; A36401; A91258; B60808; S10780; D45800; A26693; A90309; A91458; A94
R;Hollowachuk, E.W.; Stoltenberg, J.K.; Reed, R.G.; Peters Jr., T.
submitted to the EMBL Data Library, August 1991
A;Description: Bovine serum albumin: cDNA sequence and expression.
A;Reference number: A38885
A;Accession: A38885
A;Molecule type: mRNA
A;Residues: 1-607 <HO>
A;Cross-references: ENBL:M73215
R;Hiyayama, K.; Akashi, S.; Furuya, M.; Fukuhara, K.
Biochem. Biophys. Res. Commun. 173, 639-646, 1990
A;Title: Rapid confirmation and revision of the primary structure of bovine serum albumi
A;Reference number: A36401; MUID:91083649; PMID:2260975
A;Accession: A36401
A;Molecule type: protein
A;Residues: 25-41, 'H', 43-189, 'E', 191-213, 'T', 215-323, 'D', 325-393, 'TS', 396-607 <HIR>
R;MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.
Eur. J. Biochem. 98, 477-485, 1979
A;Title: Biosynthesis of bovine plasma proteins in a cell-free system.
A;Reference number: A91258; MUID:80024278; PMID:488109
A;Accession: A91258
A;Molecule type: protein
A;Residues: 1-32 <MAG>
R;Hsieh, J.C.; Lin, F.P.; Tan, M.F.
Anal. Biochem. 170, 1-8, 1988
A;Title: Electrophoretic transfer of bovine plasma proteins from an analytical isoelectrofocusing g
A;Reference number: A60808; MUID:88267456; PMID:3389500
A;Accession: B60808
A;Molecule type: protein
A;Residues: 25-41 <HS1>
R;Strawich, E.; Glimcher, M.J.
Eur. J. Biochem. 191, 47-56, 1990
A;Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelin' is albu
A;Reference number: S10780; MUID:90336641; PMID:2379503
A;Accession: S10780
A;Molecule type: protein
A;Residues: 25-41, 'H', 43-57, 59-64 <STR>
R;Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteas
A;Reference number: A45800; MUID:89341406; PMID:2474609

A;Accession: D45800
A;Molecule type: protein
A;Residues: 163-172 <CAR>
R;Carraway, R.E.; Mitra, S.P.; Cochran, D.E.
J. Biol. Chem. 262, 5968-5973, 1987
A;Title: Structure of a biologically active neurotensin-related peptide obtained from I
A;Reference number: A26693; MUID:87194805; PMID:2437111
A;Accession: A26693
A;Molecule type: protein
A;Residues: 165-172, 'L', 'CA2'
R;Reed, R.G.; Putnam, F.W.; Peters Jr., T.
Biochem. J. 191, 867-868, 1980
A;Title: Sequence of residues 400-403 of bovine serum albumin.
A;Reference number: A90309; MUID:82023364; PMID:7283978
A;Accession: A90309
A;Molecule type: protein
A;Residues: 402-433 <REE>
R;Brown, J.R.
Fed. Proc. 34, 591, 1975
A;Title: Structure of bovine serum albumin.
A;Reference number: A91458
A;Accession: A91458
A;Molecule type: protein
A;Residues: 35-41, 'H', 43-117, 'EQ', 120-179, 181-189, 'E', 191-194, 'A', 196-213, 'T', 215-288, '
R;Brown, J.R.
submitted to the Atlas, April 1975
A;Reference number: A94551
A;Accession: A94551
A;Molecule type: protein
A;Residues: 190-195 <BR2>
R;Brown, J.R.
Fed. Proc. 33, 1389, 1974
A;Reference number: A91457
A;Contents: annotation; disulfide bonds
R;Warlen, R.C.; Offord, R.E.; Rose, K.
Biochem. J. 302, 907-911, 1994
A;Title: Preparation and characterization of novel substrates of insulin proteinase (E
A;Reference number: S55232; MUID:95031935; PMID:7945219
A;Accession: S55232
A;Status: preliminary
A;Molecule type: protein
A;Residues: 529-536;569-572 <WER>
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; copper binding; duplication; plasma
F;1-18/Domain: signal sequence #status experimental <SIG>
F;19-24/Domain: propeptide #status experimental <PRO>
F;25-607/Product: serum albumin #status experimental <MPT>
F;29-201/Domain: serum albumin repeat homology <SA1>
F;220-391/Domain: serum albumin repeat homology <SA2>
F;412-591/Domain: serum albumin repeat homology <SA3>
F;727/Binding site: copper (His) #status predicted
F;77-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-392,

Query Match 100.0%; Score 45; DB 1; Length 607;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
|||||
Db 501 TESLVNRRP 509

RESULT 4
ABBSH
serum albumin precursor - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text change 22-Jun-1999
C;Accession: S06936
R;Brown, W.M.; Dziegielewska, K.M.; Foreman, R.C.; Saunders, N.R.
Nucleic Acids Res. 17, 10495, 1989
A;Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.
A;Reference number: S06936; MUID:90098886; PMID:26502160
A;Accession: S06936

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 1.09695 Seconds
(without alignments)
789.208 Million cell updates/sec

Title: US-09-832-929-18_COPY_478_486

Perfect score: 45

Sequence: 1 TESLVNRRP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	45	100.0	600	2 A47391	serum albumin prec
2	45	100.0	605	1 ABPGS	serum albumin prec
3	45	100.0	607	1 ABBS	serum albumin prec
4	45	100.0	607	1 ABBS	serum albumin prec
5	45	100.0	608	2 S57632	serum albumin prec
6	45	100.0	609	1 ABHUS	serum albumin prec
7	36	80.0	472	2 C42391	Ca2+-transporting
8	36	80.0	1205	2 A35547	Ca2+-transporting
9	35	77.8	1110	2 T33877	respiratory nitrat
10	35	77.8	1261	2 G83162	hypothetical prote
11	34	75.6	1281	2 G83405	hypothetical prote
12	33	73.3	97	2 U30106	hypothetical iLK p
13	33	73.3	473	2 A42391	Ca2+-transporting
14	33	73.3	1169	2 S94356	Ca2+-transporting
15	33	73.3	1203	2 S54357	plasma membrane Ca
16	32	71.1	315	2 H84938	flagellar motor sw
17	32	71.1	607	1 ABHOS	serum albumin prec
18	31	68.9	94	2 AG1891	hypothetical prote
19	31	68.9	95	2 J43121	gibberellin respon
20	31	68.9	242	2 S43598	membrane protein -
21	31	68.9	256	2 AB3297	hypothetical prote
22	31	68.9	264	2 C70603	hypothetical prote
23	31	68.9	267	2 D72746	Bil177_F3_123 prote
24	31	68.9	326	2 D71521	probable malate de
25	31	68.9	335	2 E95249	conserved hypotet
26	31	68.9	335	2 B98114	hypothetical prote
27	31	68.9	410	2 S85515	probable arginine
28	31	68.9	422	2 A56640	CDC4 repeat unit-c
29	31	68.9	587	2 D97590	neopullulanase (EC

30 31 68.9 609 1 FPHU
31 31 68.9 609 1 FPGO
32 31 68.9 609 1 J64258
33 31 68.9 856 2 E75292
34 31 68.9 891 2 H83218
35 30 66.7 191 2 AG0249
36 30 66.7 259 2 T21196
37 30 66.7 291 2 G81916
38 30 66.7 304 2 A72596
39 30 66.7 318 2 B81886
40 30 66.7 334 2 C86813
41 30 66.7 358 2 D82485
42 30 66.7 373 2 AB0399
43 30 66.7 401 2 F90288
44 30 66.7 465 2 AH1406
45 30 66.7 465 2 AH1782

ALIGNMENTS

RESULT 1

serum albumin precursor - rhesus macaque
A47391
C:Species: Macaca mulatta (rhesus macaque)
C>Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
C:Accession: A47391
R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, I
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993
A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bili:
A:Reference number: A47391; MUID:93211571; PMID:8460152
A:Contents: B/B homozygote
A:Accession: A47391
A>Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-600 <WAT>
A:Cross-references: GB:M90463; NID:G342294; PIDN:AAA36906.1; PID:G342295
A:Experimental source: liver
A:Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBI:128281)
C:Superfamily: serum albumin; serum albumin repeat homology
F:21-194/Domain: serum albumin repeat homology <SA1>
F:213-386/Domain: serum albumin repeat homology <SA2>
F:405-584/Domain: serum albumin repeat homology <SA3>

Query Match 100.0%; Score 45; DB 2; Length 600;

Best Local Similarity 100.0%; Pred. No. 0.13; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

OY 1 TESLVNRRP 9
|||||
Db 494 TESLVNRRP 502

RESULT 2

serum albumin precursor - pig (fragment)
ABPGS
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: S01382; A61006
R:Weinstock, J.; Baldwin, G.S.
Nucleic Acids Res. 16, 5045, 1988
A:Title: Nucleotide sequence of porcine liver albumin.
A:Reference number: S01382; MUID:89016582; PMID:3174440
A:Accession: S01382
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-505 <WEI>
A:Cross-references: EMBL:X12422; NID:G1875; PIDN:CAA30970.1; PID:G833798
R:Lineback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.
J. Bone Miner. Res. 4, 235-241, 1989
A:Title: Serum albumin and its acid hydrolysis peptides dominate preparations of miner:
A:Reference number: A61006; MUID:89269769; PMID:2728927
A:Accession: A61006

```

; SEQUENCE CHARACTERISTICS:
;   LENGTH: 610 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-984-186-2

Query Match      100.0%; Score 71; DB 4; Length 610;
Best Local Similarity 100.0%; Pred. No. 8.5e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
Db 486 VLHEKTPVSDRVTK 499

RESULT 15
US-10-153-064-133
; Sequence 133, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PE556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 133
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-133

Query Match      100.0%; Score 71; DB 4; Length 651;
Best Local Similarity 100.0%; Pred. No. 9.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
Db 528 VLHEKTPVSDRVTK 541

```

Search completed: April 19, 2004, 12:05:21
Job time : 2.14589 secs

; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
 ; TITLE OF INVENTION: Protein
 ; NUMBER OF SEQUENCES: 33
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Angen Center, Patent Operations/RRC
 ; STREET: 1840 DeHavilland Drive
 ; CITY: Thousand Oaks
 ; STATE: California
 ; COUNTRY: U.S.
 ; ZIP: 91320-1789
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/04075
 ; FILING DATE:
 ; CLASSIFICATION:
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 609 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: unknown
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; PCT-US95-04075-3

Query Match 100.0%; Score 71; DB 5; Length 609;
 Best Local Similarity 100.0%; Pred. NO. 8.5e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14
 Db 486 VLHEKTPVSDRVTK 499

RESULT 13
 US-08-797-689-2
 ; Sequence 2, Application US/08797689
 ; Patent No. 5876969
 ; GENERAL INFORMATION:
 ; APPLICANT: Fleer, Reinhard
 ; APPLICANT: Fournier, Alain
 ; APPLICANT: Guittion, Jean-Dominique
 ; APPLICANT: Jung, Gerard
 ; APPLICANT: Yeh, Patrice
 ; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
 ; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
 ; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
 ; NUMBER OF SEQUENCES: 36
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Rhone-Poulenc Rorer Inc.
 ; STREET: 500 Arcola Road, 3C43
 ; CITY: Collegeville
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19426
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: Macintosh
 ; OPERATING SYSTEM: System 7.1
 ; SOFTWARE: Word 5.1 (Patentin)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/797,689
 ; FILING DATE: 31-JAN-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/256,927
 ; FILING DATE: 28-JUL-1994
 ; APPLICATION NUMBER: FR 92/01064
 ; FILING DATE: 31-JAN-1992
 ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/FR93/00085
 ; FILING DATE: 28-JAN-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Smith Ph.D., Julie K.
 ; REGISTRATION NUMBER: P-38,619
 ; REFERENCE/DOCKET NUMBER: ST92006-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (610) 454-3839
 ; TELEFAX: (610) 454-3808
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 610 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-797-689-2
 Query Match 100.0%; Score 71; DB 2; Length 610;
 Best Local Similarity 100.0%; Pred. NO. 8.5e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14
 Db 486 VLHEKTPVSDRVTK 499

RESULT 14
 US-09-984-186-2
 ; Sequence 2, Application US/09984186
 ; Patent No. 6686179
 ; GENERAL INFORMATION:
 ; APPLICANT: Fleer, Reinhard
 ; APPLICANT: Fournier, Alain
 ; APPLICANT: Guittion, Jean-Dominique
 ; APPLICANT: Jung, Gerard
 ; APPLICANT: Yeh, Patrice
 ; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
 ; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
 ; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
 ; NUMBER OF SEQUENCES: 36
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Rhone-Poulenc Rorer Inc.
 ; STREET: 500 Arcola Road, 3C43
 ; CITY: Collegeville
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19426
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: Macintosh
 ; OPERATING SYSTEM: System 7.1
 ; SOFTWARE: Word 5.1 (Patentin)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/984,186
 ; FILING DATE: 29-Oct-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/797,689
 ; FILING DATE: 31-JAN-1997
 ; APPLICATION NUMBER: US 08/256,927
 ; FILING DATE: 28-JUL-1994
 ; APPLICATION NUMBER: FR 92/01064
 ; FILING DATE: 31-JAN-1992
 ; APPLICATION NUMBER: PCT/FR93/00085
 ; FILING DATE: 28-JAN-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Smith Ph.D., Julie K.
 ; REGISTRATION NUMBER: P-38,619
 ; REFERENCE/DOCKET NUMBER: ST92006-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (610) 454-3839
 ; TELEFAX: (610) 454-3808
 ; INFORMATION FOR SEQ ID NO: 2:

APPLICANT: Bart, Kathryn A.
 APPLICANT: Brierley, Russell A.
 APPLICANT: Thill, Gregory P.
 APPLICANT: Tschopp, Juerg F.
 TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN
 TITLE OF INVENTION: PICHIA PASTORIS

NUMBER OF SEQUENCES: 19
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Scully, Scott, Murphy & Presser
 STREET: 400 Garden City Plaza
 CITY: Garden City
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 11530-0299

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA: US/08/433,037
 APPLICATION NUMBER: US/08/433,037
 FILING DATE: 03-MAY-1995
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: DiGiglio, Frank S.
 REGISTRATION NUMBER: 31,346
 REFERENCE/DOCKET NUMBER: 9108Z

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (516) 742-4343
 TELEFAX: (516) 742-4366
 TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 609 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

MOLECULE TYPE: protein
 US-08-433-037-4

Query Match 100.0%; Score 71; DB 1; Length 609;
 Best Local Similarity 100.0%; Pred. No. 8.5e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
 DB 486 VLHEKTPVSDRVTK 499

RESULT 9
 US-08-897-956A-2
 Sequence 2, Application US/08897956A
 Patent No. 6423512

GENERAL INFORMATION:
 APPLICANT: Mary Ellen Digan
 APPLICANT: Philip Lake
 APPLICANT: Hermann Gram

TITLE OF INVENTION: Fusion Polypeptides
 FILE REFERENCE: 600-7244/CPA
 CURRENT APPLICATION NUMBER: US/08/897,956A

PRIOR FILING DATE: 1997-07-21
 PRIOR APPLICATION NUMBER: 60/022,689
 PRIOR FILING DATE: 1996-07-26

NUMBER OF SEQ ID NOS: 38
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 2
 LENGTH: 609
 TYPE: PRT
 ORGANISM: Homo Sapiens

US-08-897-956A-2
 Query Match 100.0%; Score 71; DB 4; Length 609;
 Best Local Similarity 100.0%; Pred. No. 8.5e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
 DB 486 VLHEKTPVSDRVTK 499

US-09-976-594-977
 Sequence 77, Application US/09976594
 Patent No. 6673549

GENERAL INFORMATION:
 APPLICANT: Furness, Michael
 APPLICANT: Buchbinder, Jenny

TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
 FILE REFERENCE: PA-0041 US
 CURRENT APPLICATION NUMBER: US/09/976,594

PRIOR FILING DATE: 2001-10-12
 PRIOR APPLICATION NUMBER: 60/240,409
 PRIOR FILING DATE: 2000-10-12

NUMBER OF SEQ ID NOS: 1143
 SOFTWARE: PERL Program
 SEQ ID NO 977
 LENGTH: 609
 TYPE: PRT
 ORGANISM: Homo sapiens

FEATURE:
 NAME/KEY: misc feature
 OTHER INFORMATION: Incyte ID No. 6673549 088957CD1

US-09-976-594-977
 Query Match 100.0%; Score 71; DB 4; Length 609;
 Best Local Similarity 100.0%; Pred. No. 8.5e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
 DB 486 VLHEKTPVSDRVTK 499

RESULT 12
 PCT-US95-04075-3
 Sequence 3, Application PC/TUS9504075
 GENERAL INFORMATION:
 APPLICANT: AMGEN INC.

US-10-153-064-7
 Sequence 7, Application US/10153064
 Patent No. 6663485

GENERAL INFORMATION:
 APPLICANT: Bell et al.
 TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins

FILE REFERENCES: PF556
 CURRENT APPLICATION NUMBER: US/10/153,064
 CURRENT FILING DATE: 2002-05-24

PRIOR APPLICATION NUMBER: 60/293,212
 PRIOR FILING DATE: 2001-05-25
 NUMBER OF SEQ ID NOS: 137

SOFTWARE: PatentIn version 3.1
 SEQ ID NO 7
 LENGTH: 609
 TYPE: PRT
 ORGANISM: Homo Sapiens

US-10-153-064-7
 Query Match 100.0%; Score 71; DB 4; Length 609;
 Best Local Similarity 100.0%; Pred. No. 8.5e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
 DB 486 VLHEKTPVSDRVTK 499

```
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 585 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-702-572-2
Query Match      100.0%; Score 71; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 8.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VLHEKTPVSDRVTK 14
Db      462 VLHEKTPVSDRVTK 475

RESULT 5
US-08-769-746-2
; Sequence 2, Application US/08769746
; Patent No. 6274305
; GENERAL INFORMATION:
;   APPLICANT: Sonnenschein, Carlos
;   APPLICANT: Soto, Ana M.
;   TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Medlen & Carroll, LLP
;   STREET: 220 Montgomery Street, Suite 2200
;   CITY: San Francisco
;   STATE: California
;   COUNTRY: United States of America
;   ZIP: 94104
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   FILING DATE: 19-DEC-1996
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Carroll, Peter G.
;   REGISTRATION NUMBER: 32,837
;   REFERENCE/DOCKET NUMBER: MERRI-02584
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 705-8410
;   TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 585 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-769-746-2
Query Match      100.0%; Score 71; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 8.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VLHEKTPVSDRVTK 14
Db      462 VLHEKTPVSDRVTK 475

RESULT 6
US-10-153-064-5
; Sequence 5, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
;   APPLICANT: Bell et al.
;   TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
;   FILE REFERENCE: PF556
;   CURRENT APPLICATION NUMBER: US/10/153,064
;   CURRENT FILING DATE: 2002-05-24
;   PRIOR APPLICATION NUMBER: 60/293,212
;   PRIOR FILING DATE: 2001-05-25
;   NUMBER OF SEQ ID NOS: 137
;   SOFTWARE: PatentIn version 3.1
;   SEQ ID NO 5
;   LENGTH: 585
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-153-064-5
Query Match      100.0%; Score 71; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 8.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VLHEKTPVSDRVTK 14
Db      462 VLHEKTPVSDRVTK 475

RESULT 7
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
;   APPLICANT: Lichenstein, Henri
;   APPLICANT: Lyons, David
;   APPLICANT: Wurfel, Mark
;   APPLICANT: Wright, Samuel
;   TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
;   TITLE OF INVENTION: Protein
;   NUMBER OF SEQUENCES: 33
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Amgen Center, Patent Operations/RRC
;   STREET: 1840 DeHavilland Drive
;   CITY: Thousand Oaks
;   STATE: California
;   COUNTRY: U.S.
;   ZIP: 91320-1789
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   FILING DATE:
;   CLASSIFICATION: 435
;   INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 609 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: unknown
;   TOPOLOGY: unknown
;   MOLECULE TYPE: protein
US-08-222-619-3
Query Match      100.0%; Score 71; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 8.5e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VLHEKTPVSDRVTK 14
Db      486 VLHEKTPVSDRVTK 499

RESULT 8
US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828
; GENERAL INFORMATION:
;   APPLICANT: Sreekrishna, Kotikanyadan
```

MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE: Homo sapiens
FEATURE:
NAME/KEY: Region
LOCATION: 369..419 /note= "Alternative C-termini of
OTHER INFORMATION: HSA(1-n)"
FEATURE:
NAME/KEY: Region
LOCATION: 1..585
OTHER INFORMATION: /note= "Amino acid sequence of
OTHER INFORMATION: natural HSA"
US-08-153-799-14

Query Match 100.0%; Score 71; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 8.1e-05; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14
Db 462 VLHEKTPVSDRVTK 475

RESULT 2
US-08-448-196A-3
; Sequence 3, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
; TITLE OF INVENTION: RELATED PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NASA
; STREET: MARSHALL SPACE FLIGHT CENTER
; CITY: HUNTSVILLE
; STATE: ALABAMA
; COUNTRY: USA
; ZIP: 35812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,196A
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: BROAD JR., ROBERT L.
; REGISTRATION NUMBER: 18,757
; REFERENCE/DOCKET NUMBER: XX/WFS-28402-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 205-544-0021
; TELEFAX: 205-544-0258
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-448-196A-3

Query Match 100.0%; Score 71; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 8.1e-05; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14
Db 462 VLHEKTPVSDRVTK 475

RESULT 3
US-08-984-176-1
; Sequence 1, Application US/08984176
; Patent No. 5948609
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; APPLICANT: HO, JOSEPH X
; APPLICANT: RUKER, FLORIAN
; TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT
; TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER
; FILE REFERENCE: 08/984,176
; CURRENT APPLICATION NUMBER: US/08/984,176
; CURRENT FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-984-176-1

Query Match 100.0%; Score 71; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 8.1e-05; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14
Db 462 VLHEKTPVSDRVTK 475

RESULT 4
US-08-702-572-2
; Sequence 2, Application US/08702572
; Patent No. 5965386
; GENERAL INFORMATION:
; APPLICANT: Kerry-Williams, Sean M
; APPLICANT: Gilbert, Sarah C
; TITLE OF INVENTION: Yeast Strains and Modified Albumins
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Centeon L.L.C.
; STREET: 1020 First Avenue
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-1310
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,572
; FILING DATE: 11-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 95/23857
; FILING DATE: 1-MAR-1995
; APPLICATION NUMBER: GB 9404270.2
; FILING DATE: 5-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Naomi Biswas
; REGISTRATION NUMBER: 38,384
; REFERENCE/DOCKET NUMBER: CE0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610/878/4294
; TELEFAX: 610/878/4221
; INFORMATION FOR SEQ ID NO: 2:

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:40:29 ; Search time 2.14589 Seconds

(without alignments)
336.813 Million cell updates/sec

Title: US-09-832-929-18_COPY_462_475

Perfect score: 71

Sequence: 1 VLHEKTPVSDRVTK 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	71	100.0	585	1	US-08-153-799-14
2	71	100.0	585	1	US-08-448-196A-3
3	71	100.0	585	2	US-08-984-176-1
4	71	100.0	585	2	US-08-702-572-2
5	71	100.0	585	3	US-08-763-746-2
6	71	100.0	585	4	US-10-153-064-5
7	71	100.0	609	1	US-08-222-619-3
8	71	100.0	609	1	US-08-433-037-4
9	71	100.0	609	4	US-08-897-956A-2
10	71	100.0	609	4	US-10-153-064-7
11	71	100.0	609	4	US-09-976-594-977
12	71	100.0	609	5	PCT-US95-04075-3
13	71	100.0	610	2	US-08-797-689-2
14	71	100.0	610	4	US-09-984-186-2
15	71	100.0	651	4	US-10-153-064-133
16	71	100.0	652	4	US-10-153-064-96
17	71	100.0	652	4	US-10-153-064-99
18	71	100.0	652	4	US-10-153-064-105
19	71	100.0	652	4	US-10-153-064-132
20	71	100.0	653	4	US-10-153-064-131
21	71	100.0	656	4	US-10-153-064-130
22	71	100.0	660	4	US-10-153-064-90
23	71	100.0	660	4	US-10-153-064-93
24	71	100.0	668	4	US-10-153-064-102
25	71	100.0	676	4	US-10-153-064-95
26	71	100.0	676	4	US-10-153-064-98
27	71	100.0	676	4	US-10-153-064-104

28 71 100.0 676 4 US-10-153-064-127 Sequence 127, Appl
29 71 100.0 676 4 US-10-153-064-129 Sequence 129, Appl
30 71 100.0 677 4 US-10-153-064-125 Sequence 125, Appl
31 71 100.0 680 4 US-10-153-064-123 Sequence 123, Appl
32 71 100.0 684 4 US-10-153-064-92 Sequence 92, Appl
33 71 100.0 692 4 US-10-153-064-101 Sequence 101, Appl
34 71 100.0 783 1 US-08-256-938-2 Sequence 2, Appl
35 71 100.0 787 1 US-08-256-938-4 Sequence 4, Appl
36 71 100.0 787 2 US-08-797-689-16 Sequence 16, Appl
37 71 100.0 787 4 US-09-984-186-15 Sequence 16, Appl
38 71 100.0 978 4 US-08-897-956A-3 Sequence 3, Appl
39 71 100.0 1184 4 US-10-153-064-89 Sequence 89, Appl
40 64 90.1 117 1 US-08-448-196A-2 Sequence 2, Appl
41 64 90.1 582 1 US-08-134-638-1 Sequence 1, Appl
42 64 90.1 583 1 US-08-448-196A-4 Sequence 4, Appl
43 64 90.1 583 1 US-08-448-196A-6 Sequence 6, Appl
44 64 90.1 584 1 US-08-448-196A-7 Sequence 7, Appl
45 63 88.7 583 1 US-08-448-196A-5 Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-153-799-14
; Sequence 14, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

PD 05-SEP-2002.
 XX
 XX 14-AUG-2001; 2001US-00929552.
 PF
 XX
 PR 19-DEC-1996; 96US-00769746.
 XX
 XX (TUFT) UNIV TUFTS.
 PA
 XX
 XX Sonnenschein C, Soto AM;
 PI
 XX
 XX WPI; 2003-066789/06.
 DR
 DR N-PSDB; ABX13582.
 XX
 XX Testing human cancer cells, especially breast and prostate cancer cells,
 PT by contacting cancer cells obtained from biopsy of a patient ex vivo with
 PT antibody specific to human albumin receptor.
 XX
 XX Disclosure; Fig 2; 17pp; English.
 PS
 XX The invention relates to testing human cancer cells, comprising obtaining
 CC cancer cells from the patient and contacting the cell ex vivo with an
 CC antibody to the receptor for human albumin. The method is useful for
 CC testing human cancer cells in particular breast and prostate cancer
 CC cells. The present sequence is mature human serum albumin, HSA. The anti-
 CC proliferative effect of HSA was assayed in an experiment included in the
 CC specification
 CC
 XX Sequence 584 AA;
 SQ
 Query Match 100.0%; Score 71; DB 6; Length 584;
 Best Local Similarity 100.0%; Pred. No. 0.00047;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VLHEKTPVSDRYTK 14
 Db 461 VLHEKTPVSDRYTK 474
 Search completed: April 19, 2004, 11:51:21
 Job time : 8.23453 secs

```

XX 18-DEC-2001 (first entry)
XX Novel human secreted protein #3485.
XX Human: vaccination; gene therapy; nutritional supplement;
XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;
XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX Homo sapiens.
XX WO200179449-A2.
XX 25-OCT-2001.
XX 16-APR-2001; 2001WO-US008656.
XX 18-APR-2000; 2000US-00552929.
XX 26-JAN-2001; 2001US-00770160.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-611725/70.
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy.
XX Claim 20; Page 696; 765pp; English.
XX The invention relates to novel human secreted polypeptides. The
XX polypeptides and antibodies to the polypeptides are useful for
XX determining the presence of or predisposition to a disease associated
XX with altered levels of polypeptide. The polypeptides are also useful for
XX identifying agents (agonists and antagonists) that bind to them. Cells
XX expressing the proteins are useful for identifying a therapeutic agent
XX for use in treatment of a pathology related to aberrant expression or
XX physiological interactions of the polypeptide. Vectors comprising the
XX nucleic acids encoding the polypeptides and cells genetically engineered
XX to express them are also useful for producing the proteins. The proteins
XX are useful in genetic vaccination, testing and therapy, and can be used
XX as nutritional supplements. They may be used to increase stem cell
XX proliferation, to regulate haematopoiesis; and in bone, cartilage, tendon
XX and/or nerve tissue growth or regeneration; immune suppression and/or
XX stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
XX AAU29510-AAU33304 represent the amino acid sequences of novel human
XX secreted proteins of the invention
XX Sequence 507 AA;
XX Query Match 100.0%; Score 71; DB 4; Length 507;
XX Best Local Similarity 100.0%; Pred. No. 0.0004;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 VLHEKTPVSDRVTK 14
XX 238 VLHEKTPVSDRVTK 251
XX
XX RESULT 14
XX AAU29877
XX ID AAU29877 standard; protein; 550 AA.
XX AC AAU29877;
XX 18-DEC-2001 (first entry)
XX Novel human secreted protein #368.
XX Human: vaccination; gene therapy; nutritional supplement;
XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;
XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.

```

```

XX Homo sapiens.
XX WO200179449-A2.
XX 25-OCT-2001.
XX 16-APR-2001; 2001WO-US008656.
XX 18-APR-2000; 2000US-00552929.
XX 26-JAN-2001; 2001US-00770160.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-611725/70.
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy.
XX Claim 20; Page 206; 765pp; English.
XX The invention relates to novel human secreted polypeptides. The
XX polypeptides and antibodies to the polypeptides are useful for
XX determining the presence of or predisposition to a disease associated
XX with altered levels of polypeptide. The polypeptides are also useful for
XX identifying agents (agonists and antagonists) that bind to them. Cells
XX expressing the proteins are useful for identifying a therapeutic agent
XX for use in treatment of a pathology related to aberrant expression or
XX physiological interactions of the polypeptide. Vectors comprising the
XX nucleic acids encoding the polypeptides and cells genetically engineered
XX to express them are also useful for producing the proteins. The proteins
XX are useful in genetic vaccination, testing and therapy, and can be used
XX as nutritional supplements. They may be used to increase stem cell
XX proliferation, to regulate haematopoiesis; and in bone, cartilage, tendon
XX and/or nerve tissue growth or regeneration; immune suppression and/or
XX stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
XX AAU29510-AAU33304 represent the amino acid sequences of novel human
XX secreted proteins of the invention
XX Sequence 550 AA;
XX Query Match 100.0%; Score 71; DB 4; Length 550;
XX Best Local Similarity 100.0%; Pred. No. 0.0004;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 VLHEKTPVSDRVTK 14
XX 427 VLHEKTPVSDRVTK 440
XX
XX RESULT 15
XX ABG72381
XX ID ABG72381 standard; protein; 584 AA.
XX AC ABG72381;
XX 10-FEB-2003 (first entry)
XX Mature human serum albumin #2.
XX Human: serum albumin; HSA; cancer; cytostatic; breast cancer;
XX prostate cancer; anti-proliferative.
XX Homo sapiens.
XX Key Key Location/Qualifiers
XX Misc-difference 241..242
XX /note= "Encoded by GTCACACG"
XX US2002123080-A1.

```

CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used
 CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU33304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention
 XX
 SQ Sequence 401 AA;

Query Match 100.0%; Score 71; DB 4; Length 401;
 Best Local Similarity 100.0%; Pred. No. 0.00031;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
 DB 278 VLHEKTPVSDRVTK 291
 |||||

RESULT 11
 ID AAR14179 standard; protein; 463 AA.
 AC AAR14179;
 XX 19-DEC-1991 (first entry)
 DE Human serum albumin lacking N-terminal fragment.
 XX HSA; fusion protein; drug.
 XX Homo sapiens.
 OS JP03201987-A.
 PN 03-SEP-1991.
 PD 29-DEC-1989; 89JP-00344701.
 PF 29-DEC-1989; 89JP-00344701.
 PR (TOFU) TONEN CORP.
 PA WPI; 1991-300976/41.
 DR Human serum albumin fragment - where C-terminal of human serum albumin is
 PT lacking and which can be combined with various drugs.
 PT
 PS Claim 6; Page 1; 23pp; Japanese.
 XX This sequence corresponds to amino acids 123 to 585 of mature human serum
 XX albumin. The fragment lacking the N-terminal sequence can form part of a
 CC fusion protein, for example with drugs. (This sequence is taken from the
 CC full-length HSA sequence in EP-330451). See also AAR14178
 XX
 SQ Sequence 463 AA;

Query Match 100.0%; Score 71; DB 2; Length 463;
 Best Local Similarity 100.0%; Pred. No. 0.00036;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VLHEKTPVSDRVTK 14
 |||||

DB 340 VLHEKTPVSDRVTK 353
 RESULT 12
 ID AAU3286 standard; protein; 507 AA.
 AC AAU3286;
 XX 18-DEC-2001 (first entry)
 DE Novel human secreted protein #3777.
 XX Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX Homo sapiens.
 OS WO200179449-A2.
 PN 25-OCT-2001.
 PD 16-APR-2001; 2001WO-US008656.
 PF 18-APR-2000; 2000US-00552929.
 PR 26-JAN-2001; 2001US-00770160.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Drmanac RT;
 PI WPI; 2001-611725/70.
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy.
 PT
 PS Claim 20; Page 754; 765pp; English.
 XX The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used
 CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU33304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention
 XX
 SQ Sequence 507 AA;

Query Match 100.0%; Score 71; DB 4; Length 507;
 Best Local Similarity 100.0%; Pred. No. 0.0004;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
 DB 238 VLHEKTPVSDRVTK 251
 |||||

RESULT 13
 ID AAU32994 standard; protein; 507 AA.
 AC AAU32994;
 XX
 XX

PF 16-APR-2001; 2001WO-US008656.
 XX
 PR 18-APR-2000; 2000US-00552929.
 PR 26-JAN-2001; 2001US-00770160.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-611725/70.
 XX
 PT Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy.
 XX
 PS Claim 20; Page 705; 765pp; English.
 XX

CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used
 CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU3304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention
 XX
 SQ Sequence 245 AA;

Query Match 100.0%; Score 71; DB 4; Length 245;
 Best Local Similarity 100.0%; Pred. No. 0.00018;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14
 |||||
 Db 190 VLHEKTPVSDRVTK 203
 |||||
 RESULT 9
 AAU33074
 ID AAU33074 standard; protein; 386 AA.
 AC AAU33074;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel human secreted protein #3565.
 XX
 KW Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200179449-A2.
 XX
 PD 25-OCT-2001.
 XX
 PF 16-APR-2001; 2001WO-US008656.
 XX
 PR 18-APR-2000; 2000US-00552929.
 PR 26-JAN-2001; 2001US-00770160.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-611725/70.
 DR
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy.
 PT
 XX
 PS Claim 20; Page 704; 765pp; English.
 XX
 CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used
 CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU3304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention
 XX
 SQ Sequence 386 AA;

Query Match 100.0%; Score 71; DB 4; Length 386;
 Best Local Similarity 100.0%; Pred. No. 0.0003;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14
 |||||
 Db 321 VLHEKTPVSDRVTK 334
 |||||

RESULT 10
 AAU29876
 ID AAU29876 standard; protein; 401 AA.
 AC AAU29876;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel human secreted protein #367.

XX Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200179449-A2.
 XX
 PD 25-OCT-2001.
 XX
 PF 16-APR-2001; 2001WO-US008656.
 XX
 PR 18-APR-2000; 2000US-00552929.
 PR 26-JAN-2001; 2001US-00770160.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-611725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy.
 PT
 XX
 PS Claim 20; Page 206; 765pp; English.
 XX

```
DT 28-JUL-2000 (first entry)
XX
DE Yeast codon-biased recombinant HSA protein fragment HSA-III.
XX
KW Recombinant; human serum albumin; HSA; yeast codon bias; host cell;
XX overlapping oligonucleotide; expression vector.
XX
OS Homo sapiens.
XX Synthetic.
XX
PN CN1239103-A.
XX
PD 22-DEC-1999.
XX
PF 17-JUN-1998; 98CN-00102506.
XX
PR 17-JUN-1998; 98CN-00102506.
XX
PA (HAIJ-) HAIJI BIOENGINEERING CO LTD.
XX
PI Li S, Lu D;
XX
DR WPI; 2000-351198/31.
XX
DR N-PSDE; AAA10094.
XX
PT Process for preparing recombinant human serum albumin comprising yeast
XX biased sex codons - uses a recombinant DNA technique.
XX
PS Example 1; Fig 7; 44pp; Chinese.
XX
CC The method relates to a method of recombinantly producing human serum
CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise
CC a yeast codon bias. The complete HSA gene (AAA10091) was generated as
CC three synthetic fragments (AAA10092-A10094) joined by recombinant DNA
CC technology. Each HSA fragment was synthesised from overlapping
CC oligonucleotide fragments that were extended. This sequence represents
CC the sequence of the HSA fragment HSA-III encoded by the human gene with a
CC yeast codon bias. The invention also covers a recombinant expression
CC vector, yeast host cells carrying the recombinant expression vector and
CC the process for producing human serum albumin in the yeast host cell,
XX especially in secretory mode
XX
SQ Sequence 228 AA;
Query Match 100.0%; Score 71; DB 3; Length 228;
Best Local Similarity 100.0%; Pred. NO. 0.00017;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLHEKTPVSDRVTK 14
DB 105 VLHEKTPVSDRVTK 118
|||||
RESULT 8
AAU33081
ID AAU33081 standard; protein; 245 AA.
XX
AC AAU33081;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #3572.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-C0552929.
XX
PR 26-JAN-2001; 2001US-00770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy.
XX
PS Claim 20; Page 705; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU3304 represent the amino acid sequences of novel human
XX secreted proteins of the invention
XX
SQ Sequence 151 AA;
Query Match 100.0%; Score 71; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. NO. 0.00011;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLHEKTPVSDRVTK 14
DB 35 VLHEKTPVSDRVTK 48
|||||
RESULT 7
AAU33949
ID AAU33949 standard; protein; 228 AA.
XX
AC AAU33949;
XX
```

AAU33084
ID AAU33084 standard; protein; 133 AA.

XX AC AAU33084;
XX
XX
XX DT 18-DEC-2001 (first entry)
XX
XX DE Novel human secreted protein #3575.
XX
XX Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
XX OS Homo sapiens.

XX PN WO200179449-A2.

XX PD 25-OCT-2001.

XX PF 16-APR-2001; 2001WO-US008656.

XX PR 18-APR-2000; 2000US-00552929.

XX PR 26-JAN-2001; 2001US-00770160.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI; 2001-611725/70.

XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy.
XX PS Claim 20; Page 705; 765pp; English.

XX CC The invention relates to novel human secreted polypeptides. The
XX polypeptides and antibodies to the polypeptides are useful for
XX determining the presence of or predisposition to a disease associated
XX with altered levels of polypeptide. The polypeptides are also useful for
XX identifying agents (agonists and antagonists) that bind to them. Cells
XX expressing the proteins are useful for identifying a therapeutic agent
XX for use in treatment of a pathology related to aberrant expression or
XX physiological interactions of the polypeptide. Vectors comprising the
XX nucleic acids encoding the polypeptides and cells genetically engineered
XX to express them are also useful for producing the proteins. The proteins
XX are useful in genetic vaccination, testing and therapy, and can be used
XX as nutritional supplements. They may be used to increase stem cell
XX proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
XX and/or nerve tissue growth or regeneration; immune suppression and/or
XX stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
XX AAU29510-AAU3304 represent the amino acid sequences of novel human
XX secreted proteins of the invention

XX SQ Sequence 133 AA;

Query Match 100.0%; Score 71; DB 4; Length 133;
Best Local Similarity 100.0%; Pred. No. 9.3e-05; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14

Db 17 VLHEKTPVSDRVTK 30

RESULT 5

AAU27701

ID AAU27701 standard; protein; 140 AA.

XX AC AAU27701;

XX DT 18-DEC-2001 (first entry)

XX DE Human full-length polypeptide sequence #26.

XX Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
KW nervous system disorder; inflammatory disorder; cell differentiation;
KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
KW cytostatic; antineumatic; antiarthritic; vulnery; antiinflammatory;
KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
KW neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
XX immunostimulant; analgesic; gene therapy.

XX OS Homo sapiens.

XX PN WO200164834-A2.

XX PD 07-SEP-2001.

XX PF 26-FEB-2001; 2001WO-US004926.

XX PR 28-FEB-2000; 2000US-00515126.

XX PR 18-MAY-2000; 2000US-00577409.

XX PR 17-JUN-2000; 2000US-00597707.

XX PR 14-JUL-2000; 2000US-00616807.

XX PR 19-SEP-2000; 2000US-00664641.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

XX PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;

XX PI Drmanac R;

XX DR WPI; 2001-589862/66.

XX DR N-PSDB; AA844601.

XX PT Novel polypeptides and nucleic acids obtained from cDNA libraries
XX prepared from various human tissues, for diagnosis, treatment of cancer,
XX neurological, inflammatory disorders and for use in arrays for detection.
XX Claim 10; SEQ ID NO 198; 153pp; English.

XX CC Sequences AAU27676-AAU28019 represent full-length polypeptides and contig
XX polypeptides of the invention. The proteins and their associated DNA
XX sequences are useful for the treatment, diagnosis and prevention of
XX various types of disorder in a mammalian subject such as a human, dog,
XX monkey, mouse, hamster or rat. The disorders include cancers such as
XX leukaemia, lymphoma and neuroblastoma, autoimmune disorders such as
XX multiple sclerosis, connective tissue disease, rheumatoid arthritis,
XX diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
XX disorders such as parkinson's disease, Alzheimer's disease, Huntington's
XX chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
XX Wernicke disease, inflammatory disorders such as nephritis, Crohn's
XX disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
XX bowel disease. The sequences exhibit activity relating to angiogenesis,
XX cell proliferation, cell differentiation, stem cell growth factor,
XX activin or inhibin. Therefore, they can be used to manipulate stem cells
XX in culture to give rise to neuroepithelial cells that can be used to
XX augment or replace cells damaged by illness, accidental damage or genetic
XX disorders. The sequences may also be used for regeneration of bone,
XX cartilage, tendons and ligaments and in tissue repair and burn healing.
XX Note: Some sequences for this patent did not form part of the printed
XX specification, but were obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 140 AA;

Query Match 100.0%; Score 71; DB 4; Length 140;

Best Local Similarity 100.0%; Pred. No. 9.8e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14

Db 17 VLHEKTPVSDRVTK 30

CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU3304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
XX
SQ Sequence 86 AA;

Query Match 100.0%; Score 71; DB 4; Length 86;
Best Local Similarity 100.0%; Pred. No. 5.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14
| | | | | | | | | | | | | | | |
Db 17 VLHEKTPVSDRVTK 30

RESULT 2
AAO04431
ID AAO04431 standard; protein; 106 AA.

XX AC AAO04431;

XX DT 06-NOV-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 18323.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorders; arthritis; inflammation.

XX OS Homo sapiens.

XX PN WO200164835-A2.

XX PD 07-SEP-2001.

XX PF 26-FEB-2001; 2001WO-US004927.

XX PR 28-FEB-2000; 2000US-00515126.

XX PR 18-MAY-2000; 2000US-00577409.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI; 2001-514838/56.

XX DR N-PSDB; AA184362.

XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX PT and treating e.g. leukemia, inflammation and immune disorders.

XX PS Claim 20; SEQ ID NO 18323; 1399pp + Sequence Listing; English.

XX CC The invention relates to human polynucleotides (AA179941-AA193841) and
XX CC the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The vaccines or
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation. Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 106 AA;

Query Match 100.0%; Score 71; DB 4; Length 106;
Best Local Similarity 100.0%; Pred. No. 7.2e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VLHEKTPVSDRVTK 14
| | | | | | | | | | | | | | | |
Db 88 VLHEKTPVSDRVTK 101

RESULT 3

AAU33086
ID AAU33086 standard; protein; 126 AA.

XX AC AAU33086;

XX DT 18-DEC-2001 (first entry)

XX DE Novel human secreted protein #3577.

XX KW Human; vaccination; gene therapy; nutritional supplement;
XX KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
XX KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX OS Homo sapiens.

XX PN WO200179449-A2.

XX PD 25-OCT-2001.

XX PF 16-APR-2001; 2001WO-US008656.

XX PR 18-APR-2000; 2000US-00552929.

XX PR 26-JAN-2001; 2001US-00770160.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI; 2001-611725/70.

XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic
XX PT vaccination, testing and therapy.

XX PS Claim 20; Page 705-706; 765pp; English.

XX CC The invention relates to novel human secreted polypeptides. The
XX CC polypeptides and antibodies to the polypeptides are useful for
XX CC determining the presence of or predisposition to a disease associated
XX CC with altered levels of polypeptide. The polypeptides are also useful for
XX CC identifying agents (agonists and antagonists) that bind to them. Cells
XX CC expressing the proteins are useful for identifying a therapeutic agent
XX CC for use in treatment of a pathology related to aberrant expression or
XX CC physiological interactions of the polypeptide. Vectors comprising the
XX CC nucleic acids encoding the polypeptides and cells genetically engineered
XX CC to express them are also useful for producing the proteins. The proteins
XX CC are useful in genetic vaccination, testing and therapy, and can be used
XX CC as nutritional supplements. They may be used to increase stem cell
XX CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
XX CC and/or nerve tissue growth or regeneration; immune suppression and/or
XX CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
XX CC AAU29510-AAU3304 represent the amino acid sequences of novel human
XX CC secreted proteins of the invention

XX SQ Sequence 126 AA;

Query Match 100.0%; Score 71; DB 4; Length 126;

Best Local Similarity 100.0%; Pred. No. 8.7e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14
| | | | | | | | | | | | | | | |

Db 22 VLHEKTPVSDRVTK 35

RESULT 4

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:24:29 ; Search time 8.23453 Seconds

(without alignments)
480.375 Million cell updates/sec

Title: US-09-832-929-18_COPY_462_475

Perfect score: 71

Sequence: 1 VLHEKTPVSDRVTK 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: geneseqp1980s.*
- 2: geneseqp1990s.*
- 3: geneseqp2000s.*
- 4: geneseqp2001s.*
- 5: geneseqp2002s.*
- 6: geneseqp2003as.*
- 7: geneseqp2003bs.*
- 8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	71	100.0	86	AAU33080	Novel hum
2	71	100.0	106	AAU33086	Novel hum
3	71	100.0	126	AAU33086	Novel hum
4	71	100.0	133	AAU33084	Novel hum
5	71	100.0	140	AAU27701	Human ful
6	71	100.0	151	AAU33085	Novel hum
7	71	100.0	228	AAU33081	Novel hum
8	71	100.0	245	AAU33081	Novel hum
9	71	100.0	386	AAU33074	Novel hum
10	71	100.0	401	AAU29876	Novel hum
11	71	100.0	463	AAU14179	Human ser
12	71	100.0	507	AAU33286	Novel hum
13	71	100.0	507	AAU33294	Novel hum
14	71	100.0	550	AAU29877	Novel hum
15	71	100.0	584	ABG72381	Mature hu
16	71	100.0	585	AAU33344	Sequence
17	71	100.0	585	AAU33344	Mature hu
18	71	100.0	585	AAU33344	Mature hu
19	71	100.0	585	AAU33344	Mature hu
20	71	100.0	585	AAU33344	Mature hu
21	71	100.0	585	AAU33344	Mature hu
22	71	100.0	585	AAU33344	Mature hu
23	71	100.0	585	AAU33344	Mature hu
24	71	100.0	585	AAU33344	Mature hu
25	71	100.0	585	AAU33344	Mature hu

26	71	100.0	585	2	AAW59841	Mature pr
27	71	100.0	585	3	AAW59841	Mature pr
28	71	100.0	585	3	AAW59841	Mature pr
29	71	100.0	585	4	AAW59841	Mature pr
30	71	100.0	585	4	AAW59841	Mature pr
31	71	100.0	585	4	AAW59841	Mature pr
32	71	100.0	585	4	AAW59841	Mature pr
33	71	100.0	585	4	AAW59841	Mature pr
34	71	100.0	585	4	AAW59841	Mature pr
35	71	100.0	585	4	AAW59841	Mature pr
36	71	100.0	585	4	AAW59841	Mature pr
37	71	100.0	585	4	AAW59841	Mature pr
38	71	100.0	585	5	AAW59841	Mature pr
39	71	100.0	585	5	AAW59841	Mature pr
40	71	100.0	585	5	AAW59841	Mature pr
41	71	100.0	585	5	AAW59841	Mature pr
42	71	100.0	585	5	AAW59841	Mature pr
43	71	100.0	585	5	AAW59841	Mature pr
44	71	100.0	585	6	AAW59841	Mature pr
45	71	100.0	585	6	AAW59841	Mature pr

ALIGNMENTS

RESULT 1
AAU33080
ID AAU33080 standard; protein; 86 AA.
XX
AC AAU33080;
XX
DT 18-DEC-2001 (first entry)
DE
DE Novel human secreted protein #3571.
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
FN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-00552929.
PR 26-JAN-2001; 2001US-00770160.
XX (HYSE-) HYSEQ INC.
PI Tang YT, Liu C, Drmanac RT;
XX WPI, 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
XX
PS Claim 20; Page 705; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell

```
Query Match      57.7%; Score 41; DB 16; Length 231;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVK 14
   :|||:|||||
Db 202 IIHKWGASDRIQK 215

RESULT 13
Q7VIY5 PRELIMINARY; PRT; 735 AA.
AC Q7VIY5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Flagellar basal body protein FlhA.
GN FLHA OR H0457.
OS Helicobacter hepaticus.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=32025;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 51449 / 3B1;
RX MEDLINE=22709201; PubMed=12810954;
RA Suerbaum S., Josenhans C., Stenzenbach T., Drescher B., Brandt P.,
RA Bell M., Drege M., Fartmann B., Fischer H.-P., Ge Z., Hoerster A.,
RA Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,
RA Schauer D.B., Shen Z., Weber J., Froesch M., Fox J.G.;
RT "The complete genome sequence of the carcinogenic bacterium
RT Helicobacter hepaticus."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).
RW EMBL: AE017145; AAF77064.1; -.
KW Complete proteome.
SQ SEQUENCE 735 AA; 81362 MW; 57C8AF569B377B97 CRC64;

Query Match      57.7%; Score 41; DB 16; Length 735;
Best Local Similarity 53.8%; Pred. No. 99;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVK 13
   :|||:|||||
Db 580 LLHEKVPFKDMPT 592

RESULT 14
Q8VUT7 PRELIMINARY; PRT; 358 AA.
AC Q8VUT7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phenol hydroxylase subunit PhkF.
GN PHKF.
OS Burkholderia kururiensis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=89498;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang H., Luo H., Kamagata Y.;
RT "Involvement of a phenol hydroxylase of Burkholderia kururiensis KP23
RT in TCE degradation."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RW EMBL: AB063252; BAB79284.1; -.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR006058; 2Fe2S_FeREDOXIN; 1.
DR InterPro; IPR008333; FAD_binding_6.
DR InterPro; IPR001041; Ferrredoxin.
```

```
DR InterPro; IPR001433; Oxred_FAD/NAD(P).
DR InterPro; IPR001221; Phe hydroxylase.
DR Pfam; PF00970; FAD_binding_6; 1.
DR Pfam; PF00111; fer2; 1.
DR Pfam; PF00175; NAD binding 1; 1.
DR PRINTS; PR00410; PHEHYDRYLASE.
DR PROSITE; PS00197; 2FE2S_FERREDOXIN; 1.
KW Iron; Iron-sulfur.
SQ SEQUENCE 358 AA; 38721 MW; ACCAAAF29B142BEDA CRC64;

Query Match      56.3%; Score 40; DB 2; Length 358;
Best Local Similarity 66.7%; Pred. No. 74;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LHEKTPVSDRVK 13
   :|||:|||||
Db 182 LHEQLAVGDRVT 193

RESULT 15
Q97YJ0 PRELIMINARY; PRT; 604 AA.
AC Q97YJ0;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein SS01331.
GN SS01331.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aways M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Ezauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozexa C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duquet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Senses C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
RW EMBL: AE006747; AAK41569.1; -.
DR PIR; B90289; B90289.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 604 AA; 69867 MW; CB733E807EC80B6A CRC64;

Query Match      56.3%; Score 40; DB 17; Length 604;
Best Local Similarity 61.5%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVK 13
   :|||:|||||
Db 6 ILHETTQVXDLVT 18

Search completed: April 19, 2004, 12:00:15
Job time : 7.40351 secs
```


Sequence 2, Application US/09929552
Patent No. US20020123080A1
GENERAL INFORMATION:
APPLICANT: Sonenschein, Carlos
Soto, Ana M.
TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/929,552
FILING DATE: 14-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/769,746
FILING DATE: 19-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: MERI-02584
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-929-552-2

Query Match 100.0%; Score 71; DB 9; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
DB 462 VLHEKTPVSDRVTK 475

RESULT 3
US-09-932-613-445
Sequence 445, Application US/09932613
Publication No. US20030091565A1
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
APPLICANT: Beltzer, James P.
APPLICANT: Potter, M. Daniel
APPLICANT: Fleming, Tony J.
APPLICANT: Rosen, Craig A.
TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON
FILE REFERENCE: DXX-025.1 PCT; DXX-025.1 US
CURRENT APPLICATION NUMBER: US/09/932,613
CURRENT FILING DATE: 2001-08-17
NUMBER OF SEQ ID NOS: 458
SOFTWARE: PatentIn version 3.1
SEQ ID NO 445
LENGTH: 585
TYPE: PRT
ORGANISM: Homo sapiens
US-09-932-613-445

Query Match 100.0%; Score 71; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
DB 462 VLHEKTPVSDRVTK 475

RESULT 4
US-09-984-010-26
Sequence 26, Application US/09984010
Publication No. US20030104578A1
GENERAL INFORMATION:
APPLICANT: Ballance, David James
TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
STREET: 1300 I Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA: US/09/984,010
FILING DATE: 21-May-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/091,873
FILING DATE: 25-JUN-1998
APPLICATION NUMBER: PCT/GB96/03164
FILING DATE: 19-DEC-1996
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-984-010-26

Query Match 100.0%; Score 71; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
DB 462 VLHEKTPVSDRVTK 475

RESULT 5
US-09-833-041-18
Sequence 18, Application US/09833041
Publication No. US20030125247A1
GENERAL INFORMATION:
APPLICANT: Rosen, Craig A.
APPLICANT: Haseltine, William A.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PFS45
CURRENT APPLICATION NUMBER: US/09/833,041
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/229,358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256,931

; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-041-18

Query Match 100.0%; Score 71; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
Db 462 VLHEKTPVSDRVTK 475
|||||
RESULT 6
US-09-833-117-18
; Sequence 18, Application US/09833117
; Publication No. US20030171267A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF543
; CURRENT APPLICATION NUMBER: US/09/833,117
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-117-18

Query Match 100.0%; Score 71; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
Db 462 VLHEKTPVSDRVTK 475
|||||
RESULT 7
US-09-932-322-445
; Sequence 445, Application US/09932322
; Publication No. US20030194743A1
; GENERAL INFORMATION:
; APPLICANT: Dyax Corp.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Ladner, Robert Charles
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BLyS)
; FILE REFERENCE: DYX-018.1 PCT: DYX-018.1 US
; CURRENT APPLICATION NUMBER: US/09/932,322
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445

; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-932-322-445

Query Match 100.0%; Score 71; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
Db 462 VLHEKTPVSDRVTK 475
|||||
RESULT 8
US-09-832-501-18
; Sequence 18, Application US/09832501
; Publication No. US20030199043A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J.
; APPLICANT: Sleep, Darrell
; APPLICANT: Turner, Andrew J.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF542
; CURRENT APPLICATION NUMBER: US/09/832,501
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-832-501-18

Query Match 100.0%; Score 71; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
Db 462 VLHEKTPVSDRVTK 475
|||||

US-09-832-501-18
; Sequence 18, Application US/09833118
; Publication No. US20030219875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF544
; CURRENT APPLICATION NUMBER: US/09/833,118
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT

```
; ORGANISM: Homo Sapiens
US-09-833-118-18

Query Match      100.0%; Score 71; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VLHEKTPVSDRVTK 14
      |||||
Db      462 VLHEKTPVSDRVTK 475

RESULT 10
US-09-833-245-18
; Sequence 18, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-245-18

Query Match      100.0%; Score 71; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VLHEKTPVSDRVTK 14
      |||||
Db      462 VLHEKTPVSDRVTK 475

RESULT 11
US-10-424-999-11
; Sequence 11, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-11

Query Match      100.0%; Score 71; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VLHEKTPVSDRVTK 14
      |||||
Db      462 VLHEKTPVSDRVTK 475

RESULT 12
US-10-425-000-31
; Sequence 31, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
; TITLE OF INVENTION: Angiogenesis
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-31

Query Match      100.0%; Score 71; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VLHEKTPVSDRVTK 14
      |||||
Db      462 VLHEKTPVSDRVTK 475

RESULT 13
US-10-433-108-34
; Sequence 34, Application US/10433108
; Publication No. US20040053370A1
; GENERAL INFORMATION:
; APPLICANT: Eli Lilly and Company
; TITLE OF INVENTION: GLP-1 FUSION PROTEINS
; FILE REFERENCE: X-13991
; CURRENT APPLICATION NUMBER: US/10/433,108
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: US 60/251,954
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-433-108-34

Query Match      100.0%; Score 71; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VLHEKTPVSDRVTK 14
      |||||
Db      462 VLHEKTPVSDRVTK 475

RESULT 14
US-10-153-064-5
; Sequence 5, Application US/10153064
; Publication No. US20020142814A1
```

;/ GENERAL INFORMATION:
;/ APPLICANT: Bell et al.
;/ TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
;/ FILE REFERENCE: PF556
;/ CURRENT APPLICATION NUMBER: US/10/153,064
;/ CURRENT FILING DATE: 2002-05-24
;/ PRIOR APPLICATION NUMBER: 60/293,212
;/ PRIOR FILING DATE: 2001-05-25
;/ NUMBER OF SEQ ID NOS: 137
;/ SOFTWARE: PatentIn version 3.1
;/ SEQ ID NO 5
;/ LENGTH: 585
;/ TYPE: PRT
;/ ORGANISM: Homo Sapiens
US-10-153-064-5

Query Match 100.0%; Score 71; DB 13; Length 585;
Best Local Similarity 100.0%; Pred.No. 0.00036;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VLHEKTPVSDRVTK 14
Db 462 VLHEKTPVSDRVTK 475

RESULT 15
US-10-153-604A-5
;/ Sequence 5, Application US/10153604A
;/ Publication No. US20030143191A1
;/ GENERAL INFORMATION:
;/ APPLICANT: Bell et al.
;/ FILE REFERENCE: PF556
;/ CURRENT APPLICATION NUMBER: US/10/153,604A
;/ CURRENT FILING DATE: 2002-05-24
;/ PRIOR APPLICATION NUMBER: 60/293,212
;/ PRIOR FILING DATE: 2001-05-25
;/ NUMBER OF SEQ ID NOS: 137
;/ SOFTWARE: PatentIn version 3.1
;/ SEQ ID NO 5
;/ LENGTH: 585
;/ TYPE: PRT
;/ ORGANISM: Homo Sapiens
US-10-153-604A-5

Query Match 100.0%; Score 71; DB 14; Length 585;
Best Local Similarity 100.0%; Pred.No. 0.00036;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VLHEKTPVSDRVTK 14
Db 462 VLHEKTPVSDRVTK 475

Search completed: April 19, 2004, 12:54:59
Job time : 5.89474 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:40:29 ; Search time 1.07295 Seconds
(without alignments)
336.813 Million cell updates/sec

Title: US-09-832-929-18_COPY_362_368

Perfect score: 38
Sequence: 1 AAADPHE 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pdp:*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pdp:*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pdp:*
4: /cgn2_6/ptodata/2/iaa/8B COMB.pdp:*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pdp:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pdp:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	100.0	585	1	US-08-153-799-14
2	38	100.0	585	2	US-08-702-572-2
3	38	100.0	585	3	US-08-769-746-2
4	38	100.0	585	4	US-10-153-064-5
5	38	100.0	609	1	US-08-222-619-3
6	38	100.0	609	1	US-08-433-037-4
7	38	100.0	609	4	US-08-697-956A-2
8	38	100.0	609	4	US-10-153-064-7
9	38	100.0	609	4	US-09-976-594-977
10	38	100.0	609	5	PCT-US95-04075-3
11	38	100.0	610	2	US-08-797-689-2
12	38	100.0	610	4	US-09-984-186-2
13	38	100.0	651	4	US-10-153-064-133
14	38	100.0	652	4	US-10-153-064-96
15	38	100.0	652	4	US-10-153-064-99
16	38	100.0	652	4	US-10-153-064-105
17	38	100.0	652	4	US-10-153-064-132
18	38	100.0	653	4	US-10-153-064-131
19	38	100.0	656	4	US-10-153-064-130
20	38	100.0	660	4	US-10-153-064-90
21	38	100.0	660	4	US-10-153-064-93
22	38	100.0	668	4	US-10-153-064-102
23	38	100.0	676	4	US-10-153-064-95
24	38	100.0	676	4	US-10-153-064-98
25	38	100.0	676	4	US-10-153-064-104
26	38	100.0	676	4	US-10-153-064-127
27	38	100.0	676	4	US-10-153-064-129

28 38 100.0 677 4 US-10-153-064-125
29 38 100.0 680 4 US-10-153-064-123
30 38 100.0 684 4 US-10-153-064-92
31 38 100.0 692 4 US-10-153-064-101
32 38 100.0 783 1 US-08-256-938-2
33 38 100.0 783 4 US-08-252-931A-32103
34 38 100.0 787 1 US-08-256-938-4
35 38 100.0 787 2 US-08-797-689-16
36 38 100.0 787 4 US-08-984-186-16
37 38 100.0 978 4 US-08-897-956A-3
38 38 100.0 1184 4 US-10-153-064-89
39 34 89.5 1038 4 US-08-081-385-151
40 33 86.8 409 4 US-08-489-039A-8683
41 33 86.8 409 4 US-09-252-991A-18707
42 33 86.8 506 4 US-09-252-991A-20617
43 33 86.8 1073 4 US-08-252-991A-27341
44 32 84.2 585 1 US-08-448-196A-3
45 32 84.2 585 2 US-08-984-176-1

ALIGNMENTS

RESULT 1
US-08-153-799-14
; Sequence 14, Application US/08153799
; Patent No. 576683
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

Sequence 125, App
Sequence 123, App
Sequence 92, Appl
Sequence 101, App
Sequence 2, Appli
Sequence 32103, A
Sequence 4, Appli
Sequence 16, Appl
Sequence 16, Appl
Sequence 3, Appli
Sequence 89, Appl
Sequence 151, App
Sequence 8683, App
Sequence 18707, A
Sequence 20617, A
Sequence 27341, A
Sequence 3, Appli
Sequence 1, Appli


```
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Region
LOCATION: 369-419
OTHER INFORMATION: /note= "Alternative C-termini of
OTHER INFORMATION: HSA(1-n)"
FEATURE:
NAME/KEY: Region
LOCATION: 1-585
OTHER INFORMATION: /note= "Amino acid sequence of
OTHER INFORMATION: natural HSA"
US-08-153-799-14
Query Match 100.0%; Score 38; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAADPHE 7
Db 362 AAADPHE 368
RESULT 2
US-08-702-572-2
Sequence 2, Application US/08702572
Patent No. 5965386
GENERAL INFORMATION:
APPLICANT: Kerry-Williams, Sean M
APPLICANT: Gilbert, Sarah C
TITLE OF INVENTION: Yeast Strains and Modified Albumins
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Centeon L.L.C.
STREET: 1020 First Avenue
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-1310
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,572
FILING DATE: 11-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 95/23857
FILING DATE: 1-MAR-1995
APPLICATION NUMBER: GB 9404270.2
FILING DATE: 5-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Naomi Biswas
REGISTRATION NUMBER: 38,384
REFERENCE/DOCKET NUMBER: CE0114 US
TELEPHONE: 610/878/4294
TELEFAX: 610/878/4221
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-702-572-2
Query Match 100.0%; Score 38; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAADPHE 7
Db 362 AAADPHE 368
RESULT 3
US-08-769-746-2
Sequence 2, Application US/08769746
Patent No. 6274305
GENERAL INFORMATION:
APPLICANT: Sonnenschein, Carlos
APPLICANT: Soto, Ana M.
TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/769,746
FILING DATE: 19-DEC-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: MBRI-02584
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-769-746-2
Query Match 100.0%; Score 38; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAADPHE 7
Db 362 AAADPHE 368
RESULT 4
US-10-153-064-5
Sequence 5, Application US/10153064
Patent No. 6663485
GENERAL INFORMATION:
APPLICANT: Bell et al.
TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
FILE REFERENCE: PF556
CURRENT APPLICATION NUMBER: US/10/153,064
CURRENT FILING DATE: 2002-05-24
PRIOR APPLICATION NUMBER: 60/293,212
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 137
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 585
TYPE: PRT
ORGANISM: Homo Sapiens
```

US-10-153-064-5
Query Match 100.0%; Score 38; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
DB 362 AAADPHE 368
|||||
RESULT 5
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri
; APPLICANT: Lyons, David
; APPLICANT: Wurfel, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RR
; STREET: 1840 Beavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,619
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-222-619-3

Query Match 100.0%; Score 38; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
DB 386 AAADPHE 392
|||||
RESULT 6
US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828
; GENERAL INFORMATION:
; APPLICANT: Sreekrishna, Kotikanyadan
; APPLICANT: Barr, Kathryn A.
; APPLICANT: Brierley, Russell A.
; APPLICANT: Thill, Gregory P.
; APPLICANT: Tschoep, Juerg F.
; TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN
; TITLE OF INVENTION: PICHIA PASTORIS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza

CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/433,037
; FILING DATE: 03-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9108Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-433-037-4

Query Match 100.0%; Score 38; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
DB 386 AAADPHE 392
|||||
RESULT 7
US-08-897-956A-2
; Sequence 2, Application US/08897956A
; Patent No. 6423512
; GENERAL INFORMATION:
; APPLICANT: Mary Ellen Digan
; APPLICANT: Philip Lake
; APPLICANT: Hermann Gram
; TITLE OF INVENTION: Fusion Polypeptides
; FILE REFERENCE: 600-7244/CPA
; CURRENT APPLICATION NUMBER: US/08/897,956A
; CURRENT FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/022,689
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-08-897-956A-2

Query Match 100.0%; Score 38; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
DB 386 AAADPHE 392
|||||
RESULT 8
US-10-153-064-7
; Sequence 7, Application US/10153064
; Patent No. 6663485

GENERAL INFORMATION:
APPLICANT: Bell et al.
TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
FILE REFERENCE: PF556
CURRENT APPLICATION NUMBER: US/10/153,054
CURRENT FILING DATE: 2002-05-24
PRIOR APPLICATION NUMBER: 60/293,212
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 137
SOFTWARE: Patent in version 3.1
SEQ ID NO 7
LENGTH: 609
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-153-064-7

Query Match 100.0%; Score 38; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
Db 386 AAADPHE 392

RESULT 9
US-09-976-594-977
Sequence 977, Application US/09976594
Patent No. 6673549
GENERAL INFORMATION:
APPLICANT: Furness, Michael
APPLICANT: Buchbinder, Jenny
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
FILE REFERENCE: PA-0041 US
CURRENT APPLICATION NUMBER: US/09/976,594
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: 60/240,409
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 1143
SOFTWARE: PERL Program
SEQ ID NO 977
LENGTH: 609
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6673549 088957CD1
US-09-976-594-977

Query Match 100.0%; Score 38; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
Db 386 AAADPHE 392

RESULT 10
PCT-US95-04075-3
Sequence 3, Application PC/TUS9504075
GENERAL INFORMATION:
APPLICANT: AMGEN INC.
TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Center, Patent Operations/RRC
STREET: 1840 DeHavilland Drive
City: Thousand Oaks
STATE: California
COUNTRY: U.S.
ZIP: 91320-1789

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04075
FILING DATE:
CLASSIFICATION:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 609 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
PCT-US95-04075-3

Query Match 100.0%; Score 38; DB 5; Length 609;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
Db 386 AAADPHE 392

RESULT 11
US-08-797-689-2
Sequence 2, Application US/08797689
Patent No. 5876969
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
APPLICANT: Fournier, Alain
APPLICANT: Guittion, Jean-Dominique
APPLICANT: Jung, Gerard
APPLICANT: Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 2:

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
| | | | |
DB 386 AAADPHE 392

RESULT 13
US-10-153-064-133
; Sequence 133, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 133
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-133

Query Match 100.0%; Score 38; DB 4; Length 651;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
| | | | |
DB 428 AAADPHE 434

RESULT 14
US-10-153-064-96
; Sequence 96, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 96
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-96

Query Match 100.0%; Score 38; DB 4; Length 652;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
| | | | |
DB 428 AAADPHE 434

RESULT 15
US-10-153-064-99
; Sequence 99, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556

Query Match 100.0%; Score 38; DB 4; Length 610;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-797-689-2

Query Match 100.0%; Score 38; DB 2; Length 610;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
| | | | |
DB 386 AAADPHE 392

RESULT 12
US-09-984-186-2
; Sequence 2, Application US/09984186
; Patent No. 6686173
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/984,186
FILING DATE: 29-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-984-186-2

Query Match 100.0%; Score 38; DB 4; Length 610;
Best Local Similarity 100.0%; Pred. No. 6.1;

; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 99
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-99

Query Match 100.0%; Score 38; DB 4; Length 652;
Best Local Similarity 100.0%; Pred. No. 6.5; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
|||||
Db 428 AAADPHE 434

Search completed: April 19, 2004, 12:05:20
Job time : 1.07295 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 1.09695 seconds

(without alignments)
789.208 Million cell updates/sec

Title: US-09-832-929-18_COPY_439_447

Perfect score: 51

Sequence: 1 KHPEAKRMP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:**

1: PIR1:**

2: PIR2:**

3: PIR3:**

4: PIR4:**

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	609	1 ABHUS	serum albumin prec
2	40	78.4	600	2 A47391	serum albumin prec
3	39	76.5	352	2 F83582	biotin synthase PA
4	38	74.5	55	2 T35092	probable small hyd
5	38	74.5	98	2 T26046	hypothetical prote
6	37	72.5	474	2 T15511	hypothetical prote
7	37	72.5	1036	2 H96553	unknown protein, 2
8	37	72.5	1857	2 S22775	WOT1 protein - yea
9	36	70.6	287	2 G70455	hypothetical prote
10	35	68.6	20	2 I58192	glial fibrillary a
11	35	68.6	84	2 F70611	hypothetical prote
12	35	68.6	374	2 F83916	glycolate oxidase
13	35	68.6	807	2 AC2608	phenylalanyl-trNA
14	35	68.6	807	2 B97390	phenylalanyl-trNA
15	34	66.7	124	2 G91196	hypothetical prote
16	34	66.7	124	2 B86043	unknown protein en
17	34	66.7	220	2 D64316	restriction modifi
18	34	66.7	433	2 G70056	YwI protein - Bac
19	34	66.7	578	2 C64452	restriction modifi
20	34	66.7	637	2 A53125	conserved enzyme
21	34	66.7	893	2 C82443	conserved hypothet
22	33	64.7	194	2 G75461	DNA-binding respon
23	33	64.7	206	2 AG0661	pseudouridylylate sy
24	33	64.7	249	2 AH2152	hypothetical prote
25	33	64.7	323	2 C72311	ribose ABC transpo
26	33	64.7	366	2 G71067	hypothetical prote
27	33	64.7	366	2 A75077	iron (iii) abc tra
28	33	64.7	367	2 B72644	probable transketo
29	33	64.7	368	2 D64101	glutamate 5-kinase

hypothetical prote
hypothetical prote
conserved hypothet
membrane transport
hypothetical prote
hypothetical prote
calpastatin, cardi
neurofilament medi
hypothetical prote
translocase bindin
translocase bindin
SIP1 protein - yea
hypothetical prote
DNA polymerase III
hypothetical prote
probable giutathio
hypothetical prote

ALIGNMENTS

RESULT 1

ABHUS

serum albumin precursor [validated] - human

N;Alternate names: preproalbumin

N;Contains: Kinetensin

C;Species: Homo sapiens (man)

C;Date: 29-Jul-1981 #sequence_revision 31-Jan-1997 #text change 17-Mar-2000

C;Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S3

R;Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebu

Nucleic Acids Res. 9, 6103-6114, 1981

A;Title: The sequence of human serum albumin cDNA and its expression in Escherichia col

A;Reference number: A93743; MUID:82081882; PMID:6171778

A;Accession: A93743

A;Molecule type: mRNA

A;Residues: 1-419, 'K', 421-609 <LAW>

A;Cross-references: EMBL:V00495; GB:J00078; GB:L00132; GB:L00133; NID:G28591; PIDN:CAA2

R;Dugaiczky, A.; Law, S.W.; Dennison, O.E.

Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982

A;Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.

A;Reference number: A93936; MUID:82105994; PMID:6275391

A;Accession: A93936

A;Molecule type: mRNA

A;Residues: 1-120, 'G', 122-609 <DUG>

A;Cross-references: EMBL:V00494; NID:G28589; PIDN:CAA23753.1; PID:G28590

R;Urano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.

J. Biol. Chem. 261, 3244-3251, 1986

A;Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and

A;Reference number: I39427; MUID:86140099; PMID:2419329

A;Accession: I39427

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-26 <URA>

A;Cross-references: GB:M13075; NID:G178330; PIDN:AAA51688.1; PID:G553173

R;Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.

Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994

A;Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family

A;Reference number: I59286; MUID:94181575; PMID:8134387

A;Accession: I59286

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 282-290, 'KSRFDLQ' <WAT>

A;Cross-references: GB:S69192; NID:G546032; PIDN:AB30282.1; PID:G546033

R;Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam,

Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994

A;Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl

A;Reference number: I59313; MUID:94294404; PMID:8022807

A;Accession: I59313

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 589-590, 'ALPRRVKMLLQVKLP' <MAD>

A;Cross-references: GB:S70799; NID:G547231; PIDN:AB31177.1; PID:G547232

A/Note: this frame-shift variant is designated albumin Bazzano; four additional variants
R;Menava, J.; Parrilla, R.; Ayuso, M.S.
Submitted to the EMBL Data Library, March 1995
A/Reference number: G08292
A/Accession: G01747
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-120, 'G', 122-455 <MEN>
A/Cross-references: EMBL:U22961; NID:G763428; PIDN:AAA64922.1; PID:G763431
R;Ledgarwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.
Biochem. J. 308, 321-325, 1995
A/Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex2
A/Reference number: S55314; MUID:95275251; PMID:7755581
A/Accession: S55314
A/Molecule type: protein
A/Residues: 19-27 <LED>
R;Meloun, B.; Moravsek, L.; Kostka, V.
FEBS Lett. 58, 134-137, 1975
A/Title: Complete amino acid sequence of human serum albumin.
A/Reference number: A91420; MUID:76187907; PMID:1225573
A/Accession: A91420
A/Molecule type: protein
A/Residues: 25-117, 'EQ', 120-154, 'Q', 156-193, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395-
R;Roehr, U.; Spitteller, G.; Tripler, D.
Zusatz Liebigs Ann. Chem. 9, 881-884, 1988
A/Title: Isolation and structure elucidation of middle-molecular weight peptides from ur
A/Reference number: S06422
A/Note: this paper is in German, with an English abstract
A/Accession: S06422
A/Molecule type: protein
A/Residues: 25-48 <ROE>
R;Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.
Arch. Biochem. Biophys. 305, 595-599, 1993
A/Title: Mass spectrometric identification of modifications to human serum albumin treat
A/Reference number: S36882; MUID:93384321; PMID:8373198
A/Accession: S36882
A/Molecule type: protein
A/Residues: 45-67, 141-160, 311-337, 469-490, 570-581 <FIN>
R;Kausler, E.; Spitteller, G.
Biol. Chem. Hoppe-Seyler 372, 849-855, 1991
A/Title: Bruchtaeueck aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmol
A/Reference number: S17599; MUID:92126241; PMID:1772598
A/Accession: S17599
A/Molecule type: protein
A/Residues: 25-54, 354-357, 431-447 <KAU>
A/Note: 49-Leu was also found
R;Caraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A/Title: Structures of histamins-releasing peptides formed by the action of acid proteas
A/Reference number: A45800; MUID:89341406; PMID:2474609
A/Accession: A45800
A/Molecule type: protein
A/Residues: 166-173 <CAR>
R;Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; Wa
Biochem. Biophys. Res. Commun. 136, 993-998, 1986
A/Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-tre
A/Reference number: A03239; MUID:86242180; PMID:3087352
A/Accession: A03239
A/Molecule type: protein
A/Residues: 166-173, 'L' <MOG>
R;Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, S
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990
A/Title: Mutations in genetic variants of human serum albumin found in Italy.
A/Reference number: A38255; MUID:91062352; PMID:2247440
A/Accession: A38255
A/Molecule type: protein
A/Residues: 76-111 <GAL1>
A/Accession: B38255
A/Molecule type: protein
A/Residues: 82-105, 'K', 107-110 <GAL2>
A/Note: this variant is designated albumin Vibo Valentia
A/Accession: A38255
A/Molecule type: protein

A/Residues: 76-83, 'K', 85-106 <GAL3>
A/Note: this variant is designated albumin Torino
R;Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.
Eur. J. Biochem. 214, 437-444, 1993
A/Title: The structural characterization and bilirubin-binding properties of albumin He
A/Reference number: S33298; MUID:93292504; PMID:8513793
A/Accession: S33298
A/Molecule type: protein
A/Residues: 255-263, 'B', 265-281 <MIN1>
A/Note: this variant is designated albumin Herborn
R;Minchiotti, L.; Galliano, M.; Stoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta
Biochim. Biophys. Acta 1119, 239-238, 1992
A/Title: Two albumins with identical electrophoretic mobility are produced by diffe
A/Reference number: S21078; MUID:92190239; PMID:1347703
A/Accession: S21078
A/Molecule type: protein
A/Residues: 354-356, 'K', 358-378 <MIN2>
A/Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported
R;He, X.M.; Carter, D.C.
Nature 359, 209-215, 1992
A/Title: Atomic structure and chemistry of human serum albumin.
A/Reference number: A46756; MUID:92334427; PMID:1630489
A/Contents: annotation; X-ray crystallography, 2.8 angstroms
R;Brown, J.R.; Shockley, P.; Behrens, P.Q.
in The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40
in The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40
A/Reference number: A94442
A/Contents: annotation; three-dimensional structure and disulfide bonds
R;Saber, M.A.; Stockbauer, P.; Moravsek, L.; Meloun, B.
Collect. Czech. Chem. Commun. 42, 564-579, 1977
A/Title: Disulfide bonds in human serum albumin.
A/Reference number: A90930
A/Contents: annotation; disulfide bonds
R;Jacobsen, C.
Biochem. J. 171, 453-459, 1978
A/Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding
A/Reference number: A90299; MUID:78186630; PMID:656055
A/Contents: annotation; bilirubin-binding site
R;Peters, T.; Reed, R.G.
in Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjolholm, I., eds., 11-20
A/Title: Serum albumin: conformation and active sites.
A/Reference number: A94408
A/Contents: annotation; binding sites
R;Harper, M.E.; Dugaiczky, A.
Am. J. Hum. Genet. 35, 565-572, 1983
A/Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein gene
A/Reference number: A90028; MUID:83279982; PMID:6192711
A/Contents: annotation; gene position
R;Walker, J.E.
FEBS Lett. 66, 173-175, 1976
A/Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid
A/Reference number: A46755; MUID:76257808; PMID:955075
A/Contents: annotation
R;Bohney, J.P.; Fonda, M.L.; Feldhoff, R.C.
FEBS Lett. 298, 266-268, 1992
A/Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phosph
A/Reference number: A56294; MUID:92183881; PMID:1544460
A/Contents: annotation
A/Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in
ataze activity
C/Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized
lirubin, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak
C/Comment: A large number of variants of human serum albumin have been described.
C/Genetics:
A/Gene: GDB:ALB
A/Cross-references: GDB:118990; OMIM:103600
A/Map position: 4q11-4q13
C/Superfamily: serum albumin; serum albumin repeat homology
C/Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyridox
P.1-18/Domain: signal sequence #status predicted <SIG>
P.19-24/Domain: propeptide #status experimental <PRO>
P.25-609/Product: serum albumin #status experimental <MPT>
P.29-202/Domain: serum albumin repeat homology <SA1>

F;166-174/Product: Kinetin precursor - rhesus macaque
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status predicted
F;77-86,99-115,114-125,148-193,192-201,224-270,289-277,289-303,302-313,340-385,384-393,4
F;214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

Query Match 100.0%; Score 51; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
DB 463 KHPEAKRMP 471

RESULT 2
A47391
serum albumin precursor - rhesus macaque
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
C;Accession: A47391
R;Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F.
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993
A;Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bilirubin
A;Reference number: A47391; MUID:93211971; PMID:8460152
A;Contents: B/B homozygote
A;Accession: A47391
A;Status: preliminary
A;Molecule type: mRNA; protein
A;Residues: 1-600 <WAT>
A;Cross-references: GB:M90463; NID:G342294; PIDN:AAA36906.1; PID:G342295
A;Experimental source: liver
A;Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBI:P:128281)
C;Superfamily: serum albumin; serum albumin repeat homology
F;21-194/Domain: serum albumin repeat homology <SA1>
F;213-386/Domain: serum albumin repeat homology <SA2>
F;405-584/Domain: serum albumin repeat homology <SA3>

Query Match 78.4%; Score 40; DB 2; Length 600;
Best Local Similarity 88.9%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
DB 455 KHPEAKRMP 463

RESULT 3
F83582
biotin synthase PA0500 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: F83582
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; Bz
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Latbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: F83582
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-352 <STO>
A;Cross-references: GB:AE004487; GB:AE004091; NID:G9946361; PIDN:AA03889.1; GSPDB:GN001
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: bios, PA0500
C;Superfamily: biotin synthetase

Query Match 76.5%; Score 39; DB 2; Length 352;
Best Local Similarity 77.8%; Pred. No. 9;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
DB 109 KHPSAKDMP 117

RESULT 4

T35092
probable small hydrophilic protein - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C;Accession: T35092
R;Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A
submitted to the EMBL Data Library, July 1999
A;Reference number: Z21567
A;Accession: T35092
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-55 <SEE>
A;Cross-references: EMBL:AL056884; PIDN:CAB51455.1; GSPDB:GN00070; SCOEDB:SC4G6.32
A;Experimental source: strain A3(2)
C;Genetics:

A;Gene: SCOEDB:SC4G6.32

Query Match 74.5%; Score 38; DB 2; Length 55;
Best Local Similarity 66.7%; Pred. No. 2.1;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
DB 11 KHPOAERSP 19

RESULT 5

T26046
hypothetical protein W01C8.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T26046
R;Nhan, M.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid W01C8.
A;Reference number: Z20142
A;Accession: T26046
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-98 <NHA>
A;Cross-references: EMBL:U41508; PIDN:AAA82625.1; CESP:W01C8.1
C;Genetics:
A;Gene: CESP:W01C8.1
A;Introns: 54/3; 80/1

Query Match 74.5%; Score 38; DB 2; Length 98;
Best Local Similarity 75.0%; Pred. No. 3.8;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HPEAKRMP 9
DB 13 HPEKRLP 20

RESULT 6

T15511
hypothetical protein C15C7.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T15511
R;Leimbach, D.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid C15C7.
A;Reference number: Z18363
A;Accession: T15511
A;Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA
A/Residues: 1-474 <LEI>
A/Cross-references: EMBL:U41528; NID:g1109795; PID:g1109800; PIDN:AAA83156.1; CBSP:C15C7
C/Genetics:
A/Introns: 31/3; 67/2; 106/2; 166/1; 202/3; 235/3; 364/3; 410/3
Query Match 72.5%; Score 37; DB 2; Length 474;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 KHPEAKRMP 9
Db 198 KHPETKMP 206
RESULT 7
H96553
unknown protein, 27363-23366 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C/Accession: H96553
R/Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Chinn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huizar, L.
Nature 409, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: H96553
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1036 <STO>
A/Cross-references: GB:AE005173; NID:g10092359; PIDN:AAG12768.1; GSPDB:GN00141
A/Gene: F5D21.24
C/Genetics:
A/Map position: 1
Query Match 72.5%; Score 37; DB 2; Length 1036;
Best Local Similarity 75.0%; Pred. No. 65;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 KHPEAKRM 8
Db 959 KHPETKX 966
RESULT 8
S22775
MOT1 protein - yeast (Saccharomyces cerevisiae)
N/Alternate names: protein LPP4c; protein YPL082c
C/Species: Saccharomyces cerevisiae
C/Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 21-Jul-2000
C/Accession: S22775; S61106
R/Davis, J.L.; Kunitawa, R.; Thorner, J.
Mol. Cell. Biol. 12, 1879-1892, 1992
A/Title: A presumptive helicase (MOT1 gene product) affects gene expression and is required for normal growth of yeast.
A/Reference number: S22775; MUID:92195335; PMID:1312673
A/Accession: S22775
A/Molecule type: DNA
A/Residues: 1-1867 <DAV>
A/Cross-references: EMBL:M83224; NID:g171964; PIDN:AAA34786.1; PID:g171965
R/Hall, J.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Storms, R.K.; Vo, D.H.; Wang, submitted to the EMBL Data Library, August 1995
A/Description: The sequence of Saccharomyces cerevisiae chromosome XVI left arm.
A/Reference number: S59677
A/Accession: S61106
A/Molecule type: DNA
A/Residues: 1-1867 <HAL>

A/Cross-references: EMBL:U41849; NID:g1147608; PID:g1147612; MIPS:YPL082C
C/Genetics:
A/Cross-references: SGD:S0006003; MIPS:YPL082C
A/Map position: 16L
C/Keywords: DNA binding; nucleus; transmembrane protein
F/700-716/Domain: transmembrane #status predicted <TM1>
F/1038-1054/Domain: transmembrane #status predicted <TM2>
F/1186-1202/Domain: transmembrane #status predicted <TM3>
Query Match 72.5%; Score 37; DB 2; Length 1867;
Best Local Similarity 55.6%; Pred. No. 1-2e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 KHPEAKRMP 9
Db 775 KHPEGEKLP 783
RESULT 9
G70455
hypothetical protein aq1807 - Aquifex aeolicus
C/Species: Aquifex aeolicus
C/Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C/Accession: G70455
R/Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.
Nature 392, 353-358, 1998
A/Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A/Reference number: A70300; MUID:98196666; PMID:9537320
A/Accession: G70455
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-267 <AQF>
A/Cross-references: GB:AE000757; NID:g2984092; PIDN:AAC07643.1; PID:g2984105; GB:AE0006
A/Experimental source: strain VF5
C/Genetics:
A/Gene: aq1807
Query Match 70.6%; Score 36; DB 2; Length 267;
Best Local Similarity 75.0%; Pred. No. 25;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 KHPEAKRM 8
Db 138 KHPEKRL 145
RESULT 10
I58192
glial fibrillary acidic protein GAPP - rat (fragment)
C/Species: Rattus sp. (rat)
C/Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 27-Feb-1997
C/Accession: I58192
R/Laping, N.J.; Morgan, T.E.; Nichols, N.R.; Rozovsky, I.; Young-Chan, C.S.; Zarow, C.; Neurosciences 58, 563-572, 1994
A/Title: Transforming growth factor-beta 1 induces neuronal and astrocyte genes: tubulin
A/Reference number: I58192; MUID:9424314; PMID:8170537
A/Accession: I58192
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-20 <RES>
A/Cross-references: GB:S70248; NID:G546523
A/Experimental source: strain Fisher
C/Genetics:
A/Introns: 11/2
Query Match 68.6%; Score 35; DB 2; Length 20;
Best Local Similarity 62.5%; Pred. No. 2.8;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHPEAKRM 8
Db 138 KHPEKRL 145

Db 5 RHPEARL 12

RESULT 11

D70611

Hypothetical protein Rv0623 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 03-Nov-2000

C:Accession: D70611

R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: D70611

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-84 <COL>

A:Cross-references: GB:292772; GB:AL123456; NID:93261722; PIDN:CAB07112.1; PID:e306561;

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: Rv0623

C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0608

Query Match 68.6%; Score 35; DB 2; Length 84;

Best Local Similarity 75.0%; Pred. No. 12;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KHPEAKRM 8

Db 6 KHPEADRL 13

RESULT 12

F83916

Glycolate oxidase subunit BH2134 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: F83916

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: F83916

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-374 <STO>

A:Cross-references: GB:AP001514; GB:BA000004; NID:910174613; PIDN:BA058553.1; GSPDB:GN00

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH2134

Query Match 68.6%; Score 35; DB 2; Length 374;

Best Local Similarity 85.7%; Pred. No. 56;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKR 7

Db 222 KHPEAKO 228

RESULT 13

AC2608

phenylalanyl-tRNA synthetase beta chain [imported] - Agrobacterium tumefaciens (strain C

C:Species: Agrobacterium tumefaciens

C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C:Accession: AC2608

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AC2608

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-807 <KUR>

A:Cross-references: GB:AE008698; PIDN:AA141281.1; PID:G17738589; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: phei

A:Map position: circular chromosome

C:Superfamily: phenylalanine-tRNA ligase beta chain

Query Match 68.6%; Score 35; DB 2; Length 807;

Best Local Similarity 75.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KHPEAKRM 8

Db 54 KHPEADRL 61

RESULT 14

B97390

phenylalanyl-tRNA synthetase beta chain (phenylalanine-tRNA ligase beta chain) (pherS)

C:Species: Agrobacterium tumefaciens

C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002

C:Accession: B97390

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A:Reference number: A97359; MUID:21608551; PMID:11743194

A:Accession: B97390

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-807 <KUR>

A:Cross-references: GB:AE007869; PIDN:AAK86075.1; PID:G15155152; GSPDB:GN00169

C:Genetics:

A:Gene: AGR_C_445

A:Map position: circular chromosome

C:Superfamily: phenylalanine-tRNA ligase beta chain

Query Match 68.6%; Score 35; DB 2; Length 807;

Best Local Similarity 75.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KHPEAKRM 8

Db 54 KHPEADRL 61

RESULT 15

C91196

hypothetical protein ECs4539 [imported] - Escherichia coli (strain O157:H7, substrain R

C:Species: Escherichia coli

C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001

C:Accession: C91196

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen

A:Reference number: A98629; MUID:21156231; PMID:11258796

A:Accession: C91196

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-124 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA037962.1; PID:G13364014; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RIMD 0509952

C:Genetics:

A:Gene: ECs4539

Query Match 66.7%; Score 34; DB 2; Length 124;
 Best Local Similarity 100.0%; Pred.No. 28;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHPEAK 6
 Db 119 KHPEAK 124

Search completed: April 19, 2004, 12:02:31
 Job time : 2.09695 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 0.65651 Seconds
(without alignments)
713.823 Million cell updates/sec

Title: US-09-832-929-18_COPY_439_447
Perfect score: 51
Sequence: 1 KHPKAKEMP 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	609	1	ALBU HUMAN
2	44	86.3	608	1	ALBU RABIT
3	40	78.4	600	1	ALBU_MACMU
4	37	72.5	381	1	OFUI_CABEL
5	37	72.5	1867	1	MOTI_YEAST
6	36	70.6	1902	1	SMFI_HUMAN
7	35	68.6	461	1	NIFI_NYNP8
8	35	68.6	807	1	NIYE_AGRPT5
9	34	66.7	220	1	Y132_METJA
10	34	66.7	433	1	YWPO_EACSU
11	34	66.7	578	1	YC20_METUA
12	34	66.7	637	1	T4BA_BACCO
13	33	64.7	206	1	RLUA_YERPE
14	33	64.7	368	1	PROB_HAEIN
15	33	64.7	499	1	CP8B_RABIT
16	33	64.7	713	1	ICAL_PIG
17	33	64.7	837	1	SECA_LISIN
18	33	64.7	837	1	SECA_LISMO
19	33	64.7	838	1	RIRI_TRYBB
20	33	64.7	863	1	SIP1_YEAST
21	33	64.7	1367	1	DPO3_THEMA
22	33	64.7	1539	1	SMKY_HUMAN
23	32.5	63.7	2161	1	SHK1_RAT
24	32.5	63.7	2167	1	SHK1_RAT
25	32	62.7	158	1	RS10_SPOFR
26	32	62.7	225	1	PYRH_PYRHO
27	32	62.7	231	1	PYRH_THEMA
28	32	62.7	277	1	MTD_THEMA
29	32	62.7	437	1	YGY3_HALSQ
30	32	62.7	464	1	PNTB_RHORU
31	32	62.7	519	1	RNFC_RHOCA
32	32	62.7	593	1	GRN_HUMAN
33	32	62.7	608	1	ALBU_FELCA

34 32 62.7 695 1 CIO4_HUMAN P56696 homo sapien
35 32 62.7 787 1 SYFB_NEIMA Q9JVA0 neisseria m
36 32 62.7 787 1 SYFB_NEIMB Q9K089 neisseria m
37 32 62.7 791 1 SYFB_XANCP Q9P726 xanthomonas
38 32 62.7 792 1 SYFB_XANAC Q9P765 xanthomonas
39 32 62.7 809 1 SYFB_RALSO Q8XZ24 talstonia s
40 32 62.7 1086 1 NNTM_BOVIN P11024 bos taurus
41 32 62.7 1086 1 NNTM_HUMAN Q13423 homo sapien
42 32 62.7 1086 1 NNTM_MOUSE Q81941 mus musculus
43 32 62.7 1402 1 BIRG_MOUSE Q9J153 mus musculus
44 32 62.7 1403 1 BIRA_MOUSE Q9GWK5 mus musculus
45 32 62.7 1403 1 BIRE_MOUSE Q9R016 mus musculus

ALIGNMENTS

RESULT 1
ALBU_HUMAN STANDARD; PRT; 609 AA.
ID ALBU_HUMAN STANDARD; PRT; 609 AA.
AC P02768; Q95574; Q13140; Q9P157; Q9P117; Q9UHS3; Q9UJZ0;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86196112; PubMed=3009475;
RA Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,
RA Beattie W.G., Dugaiczak A.;
RT "Molecular structure of the human albumin gene is revealed by
RT nucleotide sequence within q11-22 of chromosome 4.";
RL J. Biol. Chem. 261:6747-6757(1986).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT LYS-420.
RX MEDLINE=82081882; PubMed=6171778;
RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,
RA Najarian R.C., Seeburg P.H., Wion K.L.;
RT "The sequence of human serum albumin cDNA and its expression in E.
RT coli.";
RN Nucleic Acids Res. 9:6103-6114 (1981).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANT GLY-121.
RX MEDLINE=82105994; PubMed=6275391;
RA Dugaiczak A., Law S.W., Dennison O.E.;
RT "Nucleotide sequence and the encoded amino acids of human serum
RT albumin mRNA.";
RN Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;
RN Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (PRO903/PRO1708/PRO2044/PRO2619/PRO2675).
RC TISSUE=Fetal liver;
RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,
RA Xu W., Gao F., Liu M., He P., Zhang Y., Ouyang S., Luo L.;
RT "Functional prediction of the coding sequences of 121 new genes
RT deduced by analysis of cDNA clones from human fetal liver.";
RN Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.
RA Huang M.C., Wu H.T.;
RT "The cDNA sequences of human serum albumin.";
RN Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.

- RC TISSUE=Liver, and Skeletal muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Wagner L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshikiyuki S., Carninci P., Frange C.,
RA Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahay J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shervchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Genetic and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [8].
RP SEQUENCE OF 25-609.
RX MEDLINE=76187907; PubMed=1225573;
RA Meloun B., Moravek L., Kostka V.;
RT "Complete amino acid sequence of human serum albumin";
RL FEBS Lett. 58:134-137(1975).
RN [9].
RP SEQUENCE OF 25-609.
RX Brown J.R., Shockley P., Behrens P.O.;
RL (In) Bing D.H. (eds.);
RT The chemistry and physiology of the human plasma proteins, pp.23-40,
RL Pergamon Press, New York (1979).
RN [10].
RP SEQUENCE OF 1-455 FROM N.A.
RC TISSUE=Liver;
RA Menaya J., Parrilla R., Ayuso M.S.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [11].
RP SEQUENCE OF 1-26 FROM N.A.
RX MEDLINE=86140099; PubMed=2419329;
RA Urano Y., Watanabe K., Sakai M., Tanaoki T.;
RT "The human albumin gene. Characterization of the 5' and 3' flanking
RT regions and the polymorphic gene transcripts";
RL J. Biol. Chem. 261:3244-3251(1986).
RN [12].
RP SEQUENCE OF 222-229.
RX MEDLINE=76257808; PubMed=955075;
RA Walker J.E.;
RT "Lysine residue 199 of human serum albumin is modified by
RT acetylsalicylic acid";
RL FEBS Lett. 66:1173-1175(1976).
RN [13].
RP SEQUENCE OF 25-44 AND 480-499.
RC TISSUE=Heart;
RX MEDLINE=95203287; PubMed=7895732;
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
RT "The human myocardial two-dimensional gel protein database: update
RT 1994";
RL Electrophoresis 15:1459-1465(1994).
RN [14].
RP DISULFIDE BONDS.
RA Sabar M.A., Stockbauer P., Moravek L., Meloun B.;
RT "Disulfide bonds in human serum albumin";
RL Collect. Czech. Chem. Commun. 42:564-579(1977).
RN [15].
RP BILIRUBIN-BINDING SITE.
RX MEDLINE=78186630; PubMed=656055;
RA Jacobsen C.;
RT "Lysine residue 240 of human serum albumin is involved in high-
RT affinity binding of bilirubin";
RL Biochem. J. 171:453-459(1978).
- RN [16].
RP VARIANT CANTERBURY ASN-337.
RX MEDLINE=87157744; PubMed=3828358;
RA Brennan S.O., Herbert P.;
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second
RT domain of serum albumin";
RN Biochim. Biophys. Acta 912:191-197(1987).
RN [17].
RP VARIANTS NAG-2 AND NAG-3.
RX MEDLINE=89068523; PubMed=3479777;
RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,
RA Satoh C., Neel J.V.;
RT "Amino acid substitutions in inherited albumin variants from
RT Amerindian and Japanese populations";
RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).
RN [18].
RP VARIANTS NAG-1, HIR-1, HIR-2 AND TOCHIGI.
RX MEDLINE=89345611; PubMed=2762316;
RA Arai K., Madison J., Huse K., Ishioka N., Satoh C., Fujita M.,
RA Neel J.V., Sakurabayashi I., Putnam F.W.;
RT "Point substitutions in Japanese allolbumins";
RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).
RN [19].
RP VARIANTS MANAUS; OSAKA; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.
RX MEDLINE=90115905; PubMed=2404284;
RA Arai K., Madison J., Shimizu A., Putnam F.W.;
RT "Point substitutions in albumin genetic variants from Asia";
RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).
RN [20].
RP DESCRIPTION OF VARIANT REDHILL.
RX MEDLINE=90115852; PubMed=2104980;
RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;
RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of
RT human serum albumin whose precursor has an aberrant signal peptidase
RT cleavage site";
RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).
RN [21].
RP VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.
RX MEDLINE=91082352; PubMed=2247440;
RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,
RA Watkins S., Putnam F.W.;
RT "Mutations in genetic variants of human serum albumin found in
RT Italy";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).
RN [22].
RP VARIANT VENEZIA.
RX MEDLINE=91296740; PubMed=2068071;
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,
RA Minchiotti L., Putnam F.W.;
RT "A donor splice mutation and a single-base deletion produce two
RT carboxyl-terminal variants of human serum albumin";
RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).
RN [23].
RP VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;
RX KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.
RX MEDLINE=92052189; PubMed=1946412;
RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,
RA Matsuda Y.-I., Amaki I., Putnam F.W.;
RT "Genetic variants of serum albumin in Americans and Japanese";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).
RN [24].
RP VARIANT CASEBROOK ASN-518.
RX MEDLINE=91316157; PubMed=1859851;
RA Peach R.J., Brennan S.O.;
RT "Structural characterization of a glycoprotein variant of human serum
RT albumin: albumin Casebrook (494 Asp-->Asn)";
RL Biochim. Biophys. Acta 1097:49-54(1991).
RN [25].
RP VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.
RX MEDLINE=92190239; PubMed=1347703;
RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,
RA Rochu D., Porta F.;
RT "Two allolbumins with identical electrophoretic mobility are produced

Query Match 100.0%; Score 51; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.042; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHEPEAKRMP 9
DB 463 KHEPEAKRMP 471

RESULT 2
ALBU_RABIT STANDARD; PRT; 608 AA.
AC F49065;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=liver;
RA Sheffield W.P., Syed S., Schuyler P.D.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Plasma.
CC -1- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC -1- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U18344; AB589347.1; --
CC HSP; P02768; 1E7B.
CC InterPro; IPR000264; Serum albumin.
CC Pfam; PF00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum albumin; 1.
CC SMART; SM00103; ALBUMIN; 3.
CC PROSITE; PS00212; ALBUMIN; 3.
CC Metal-binding; Lipid-binding; Repeat; Signal; Copper.
CC SIGNAL 1 18
CC FT PROPEP 19 24 BY SIMILARITY.
CC FT CHAIN 25 608 SERUM ALBUMIN.
CC FT DOMAIN 25 205 ALBUMIN 1.
CC FT DOMAIN 212 397 ALBUMIN 2.
CC FT DOMAIN 404 595 ALBUMIN 3.
CC FT METAL 27 27 COPPER.
CC FT DISULFID 77 86 BY SIMILARITY.
CC FT DISULFID 99 115 BY SIMILARITY.
CC FT DISULFID 114 125 BY SIMILARITY.
CC FT DISULFID 148 193 BY SIMILARITY.
CC FT DISULFID 192 201 BY SIMILARITY.
CC FT DISULFID 224 270 BY SIMILARITY.
CC FT DISULFID 269 277 BY SIMILARITY.
CC FT DISULFID 289 303 BY SIMILARITY.
CC FT DISULFID 302 313 BY SIMILARITY.
CC FT DISULFID 340 385 BY SIMILARITY.
CC FT DISULFID 384 393 BY SIMILARITY.
CC FT DISULFID 416 462 BY SIMILARITY.

FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 485 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 528 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
SQ SEQUENCE 608 AA; 6914 MW; CF5S92647AAFE9A2 CRC64;

Query Match 86.3%; Score 44; DB 1; Length 608;
Best Local Similarity 77.8%; Pred. No. 0.9;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHEPEAKRMP 9
DB 463 KHEPEAKRMP 471

RESULT 3
ALBU_MACMU STANDARD; PRT; 600 AA.
AC Q28522;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor (Fragment).
GN ALB.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=93211971; PubMed=8460152;
RA Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,
RW Dwyer J., Putnam F.W.;
RL "cDNA and protein sequence of polymorphic macaque albumins that differ
RL in bilirubin binding.";
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Plasma.
CC -1- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC -1- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M90463; AAA36906.1; --
CC PIR; A47391; A47391.
CC HSP; P02768; 1E7B.
CC InterPro; IPR000264; Serum albumin.
CC Pfam; PF00273; transport prot; 3.
CC PRINTS; PR00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum albumin; 1.
CC SMART; SM00103; ALBUMIN; 3.
CC PROSITE; PS00212; ALBUMIN; 3.
CC Metal-binding; Lipid-binding; Repeat; Signal; Copper.
CC SIGNAL 1 1
CC FT PROPEP <1 10 BY SIMILARITY.
CC FT CHAIN 11 16 BY SIMILARITY.
CC FT DOMAIN 17 600 SERUM ALBUMIN.
CC FT DOMAIN 17 197 ALBUMIN 1.
CC FT DOMAIN 204 389 ALBUMIN 2.
CC FT DOMAIN 396 587 ALBUMIN 3.
CC FT METAL 19 19 COPPER (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the SNF2/RAD54 helicase family.
CC -!- SIMILARITY: Contains 6 HEAT repeats.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M83224; AAA34786.1; -.
CC EMBL; U41849; AAB68257.1; -.
CC PIR; S22775; S22775.
CC GenOnline; L44064; -.
CC TRANSFAC; T03499; -.
CC SGD; S0006003; MOT1.
CC GO; GO:0002228; C:nuclear chromosome; IDA.
CC GO; GO:0016887; F:ATPase activity; IDA.
CC GO; GO:0006357; P:regulation of transcription from Pol II pro. .; IDA.
CC InterPro; IPR008938; ARM.
CC InterPro; IPR001410; DEAD.
CC InterPro; IPR000357; HEAT.
CC InterPro; IPR001650; Helicase C.
CC InterPro; IPR000330; SNF2_N_1.
CC Pfam; PF00271; Helicase C7.1.
CC Pfam; PF00176; SNF2_N; 1.
CC SMART; SM00487; DEADC; 1.
CC SMART; SM00490; HELIC; 1.
CC PROSITE; PS50077; HEAT_REPEAT; FALSE NEG.
CC Nuclear protein; DNA-binding; Helicase; ATP-binding; Repeat.
CC FT DOMAIN 195 211 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT REPEAT 289 326 HEAT 1.
FT REPEAT 445 482 HEAT 2.
FT REPEAT 541 578 HEAT 3.
FT REPEAT 1108 1145 HEAT 4.
FT REPEAT 1188 1225 HEAT 5.
FT REPEAT 1495 1537 HEAT 6.
FT NP_BIND 1297 1304 ATP (POTENTIAL).
FT SITE 1408 1411 DEGR BOX.
FT SEQUENCE 1867 AA; 209975 MW; 1A00005148D5632B CRC64;
SQ
Query Match 72.5%; Score 37; DB 1; Length 1867;
Best Local Similarity 55.6%; Pred. No. 60;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 KHPEAKRMP 9
Db 775 KHPEGEKLP 783
|||||
|||||
RESULT 6
SMFL_HUMAN STANDARD; PRT; 1902 AA.
AC O14497; Q9UPZ1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE SWI/SNF-related, matrix-associated, actin-dependent regulator of
DE chromatin subfamily F member 1 (SWI-SNF complex protein p270) (5120).
GN SMARCF1 OR C10RF4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A. AND MUTAGENESIS.
RX MEDLINE=2021560; PubMed=1075779;
RA Dallas P.B., Pachione S., Wilsker D., Bowrin V., Kobayashi R.,
RT Moran E.;
RT "The human SWI-SNF complex protein p270 is an ARID family member with
RT non-sequence-specific DNA binding activity.";

RL Mol. Cell. Biol. 20:3137-3146(2000).
RN [2].
RP SEQUENCE OF 1-1175 FROM N.A.
RX MEDLINE=98094256; PubMed=9434167;
RA Takeuchi T., Chen B.-K., Qiu Y., Sonobe H., Ohtsuki Y.;
RT "Molecular cloning and expression of a novel human cDNA containing CAG
RT repeats.";
RL Gene 204:71-77(1997).
RN [3].
RP SEQUENCE OF 1-1132 FROM N.A.
RA Takeuchi T., Misaki A.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: BINDS DNA NON-SPECIFICALLY.
CC -!- SUBUNIT: Part of the SWI-SNF complex.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- TISSUE SPECIFICITY: Highly expressed in spleen, thymus, prostate,
CC testis, ovary, small intestine, colon, and pbl, and at a much
CC lower level in heart, brain, placenta, lung, liver, skeletal
CC muscle, kidney, and pancreas.
CC -!- SIMILARITY: Contains 1 ARID domain.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to a
CC frameshift in position 1132.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF265208; AAF75765.1; ALT INIT.
DR EMBL; AB001895; BAA3269.1; ALT FRAME.
DR EMBL; AB024075; BAA83073.1; ALT_SEQ.
DR EMBL; AB024059; BAA83073.1; JOINED.
DR EMBL; AB024060; BAA83073.1; JOINED.
DR EMBL; AB024061; BAA83073.1; JOINED.
DR EMBL; AB024062; BAA83073.1; JOINED.
DR EMBL; AB024063; BAA83073.1; JOINED.
DR EMBL; AB024064; BAA83073.1; JOINED.
DR EMBL; AB024065; BAA83073.1; JOINED.
DR EMBL; AB024066; BAA83073.1; JOINED.
DR EMBL; AB024067; BAA83073.1; JOINED.
DR EMBL; AB024068; BAA83073.1; JOINED.
DR EMBL; AB024069; BAA83073.1; JOINED.
DR EMBL; AB024070; BAA83073.1; JOINED.
DR EMBL; AB024071; BAA83073.1; JOINED.
DR EMBL; AB024072; BAA83073.1; JOINED.
DR EMBL; AB024073; BAA83073.1; JOINED.
DR EMBL; AB024074; BAA83073.1; JOINED.
DR Genew; HGNC:11110; SMARCF1.
DR XIM; 603024; -.
DR GO; GO:0005634; C:nucleus; NAS.
DR GO; GO:0003677; F:DNA binding; NAS.
DR InterPro; IPR001606; ARID.
DR InterPro; IPR008938; ARM.
DR Pfam; PF01388; ARID; 1.
DR SMART; SM00501; BRIGHT; 1.
KW DNA-binding; Nuclear protein.
FT DOMAIN 96 99 POLY-GLN.
FT DOMAIN 178 184 POLY-GLN.
FT DOMAIN 615 618 POLY-SER.
FT DOMAIN 631 740 ARID.
FT DOMAIN 944 1021 GLN-RICH.
FT MUTAGEN 690 690 W-A: PARTIAL LOSS OF DNA-BINDING
FT ACTIVITY. WHEN A.A. 713 ALSO MUTATED,
FT COMPLETE LOSS OF ACTIVITY.
FT Y-A: PARTIAL LOSS OF DNA-BINDING
FT ACTIVITY. WHEN A.A. 690 ALSO MUTATED,
FT COMPLETE LOSS OF ACTIVITY.
FT D -> G (IN REF. 1).
FT V -> M (IN REF. 1).
FT Q -> S (IN REF. 2).
FT CONFLICT 27 27
FT CONFLICT 51 51
FT CONFLICT 349 349


```
FT CONFLICT 368 G -> GG (IN REF. 3).
FT CONFLICT 374 P -> S (IN REF. 2).
FT CONFLICT 393 P -> L (IN REF. 2).
FT CONFLICT 489 MANPPQVSGNCP -> NGQYATSGWRDVS (IN REF. 2).
FT CONFLICT 634 E -> G (IN REF. 2).
FT CONFLICT 924 P -> S (IN REF. 2).
FT CONFLICT 924 Q -> L (IN REF. 2).
FT CONFLICT 1016 Q -> P (IN REF. 2).
FT CONFLICT 1033 Q -> P (IN REF. 2).
SQ SEQUENCE 1902 AA; 205946 MW; 9B70A46F81062EAA CRC64;

Query Match
Best Local Similarity 70.6%; Score 36; DB 1; Length 1902;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KHPKAKRMP 9
||| | | |
Db 1637 KHPKQAP 1645

RESULT 7
ID_NIFN SYNPS STANDARD; PRT; 461 AA.
AC Q07356; 1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 41, Last annotation update)
DE Nitrogenase iron-molybdenum cofactor biosynthesis protein nifN.
GN NIFN.
OS Synechococcus sp. (strain PCC 8901 / RF-1) (Cyanothecae PCC 8901).
OC Bacteria; Cyanobacteria; Chroococcales; Cyanothecae.
OX NCBI_TaxID=41431;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99231861; PubMed=10217509;
RA Huang T.C., Lin R.F., Chu M.K., Chen H.M.;
RA "Organization and expression of nitrogen-fixation genes in the aerobic
RA nitrogen-fixing unicellular cyanobacterium Synechococcus sp. strain
RA RF-1.";
RA Microbiology 145:743-753(1999).
RT
RL Microbiology 145:743-753(1999).
CC -1- FUNCTION: This protein may play a role in the biosynthesis of the
CC prosthetic group of nitrogenase (FeMo cofactor).
CC -1- PATHWAY: Fe-Mo cofactor biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE NIFD/NIFK/NIFE/NIFN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF003700; AAC35195.1; -.
DR HSSP; P11347; IMIO.
DR InterPro; IPR005975; NifN.
DR InterPro; IPR000318; Nitrogenase comp1.
DR InterPro; IPR000510; Oxred nitrogenase1.
DR Pfam; PF00148; oxidized_nitro; 1.
DR TIGRFAMs; TIGR01285; nifN; 1.
DR PROSITE; PS00699; NITROGENASE_1; 1.
DR PROSITE; PS00090; NITROGENASE_1_2; FALSE_NEG.
KW Nitrogen fixation.
SQ SEQUENCE 461 AA; 50377 MW; 047765AA3BBE0A2F CRC64;

Query Match
Best Local Similarity 68.6%; Score 35; DB 1; Length 461;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KHPKAKRMP 9
||| | | |
Db 121 KHPQKMLP 129
```

RESULT 8

```
SYTB AGRT5
ID SYTB AGRT5 STANDARD; PRT; 807 AA.
AC Q8UIN4;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)
DE (Phenylalanine-tRNA ligase beta chain) (PHERS).
DE PHET OR ATU0259 OR AGR_C_445
GN Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OS Agrobacterium tumefaciens; Alphaproteobacteria; Rhizobiales;
OC Bacteri; Proteobacteria; Alphaproteobacteria; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.P., Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D.S., Grant C.,
RA Chapman P., Clendinning J., Deatherage G., Gillet W., Palmeri A.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RA "The genome of the natural genetic engineer Agrobacterium tumefaciens
RA C58.";
RA Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quorillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RA "Genome sequence of the plant pathogen and biotechnology agent
RA Agrobacterium tumefaciens C58.";
RA Science 294:2323-2328(2001).
CC -1- CATALYTIC ACTIVITY: ATP + L-phenylalanyl-tRNA(Phe) = AMP +
CC diphosphate + L-phenylalanyl-tRNA(Phe).
CC -1- COFACTOR: Binds 2 magnesium ions per tetramer (By similarity).
CC -1- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the phenylalanyl-tRNA synthetase beta chain
CC family. Subfamily 1.
CC -1- SIMILARITY: Contains 1 tRNA-binding domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE008998; AAL41281.1; -.
DR EMBL; AE007965; AAK86075.1; -.
DR PIR; B97390; B97390.
DR HAMAP; MF 00283; -.
DR InterPro; IPR005146; B3_4.
DR InterPro; IPR005147; B5_
DR InterPro; IPR005121; Fdx-AntiCB.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR InterPro; IPR004532; Pher_bact.
DR InterPro; IPR002547; tRNA_bind.
DR Pfam; PF03483; B3_4; 1.
```

DR Pfam; PF03484; B5; 1.
DR Pfam; PF03147; FDX-ACB; 1.
DR Pfam; PF01588; TRNA_Bind; 1.
DR TIGRFAMs; TIGR00472; phet_bact; 1.
DR PROSITE; PS00886; TRSD; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Metal-binding; Magnesium; RNA-binding; trna-binding;
KW Complete proteome.
FT METAL 38 148 TRNA-BINDING.
FT METAL 453 453 MAGNESIUM (BY SIMILARITY).
FT METAL 459 459 MAGNESIUM (VIA CARBONYL OXYGEN) (BY
SIMILARITY).
FT METAL 462 462 MAGNESIUM (BY SIMILARITY).
FT METAL 463 463 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 807 AA; 86356 MW; OD36D17DD8F9CEB0 CRC64;
Query Match 68.6%; Score 35; DB 1; Length 807;
Best Local Similarity 75.0%; Pred. No. 61;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 KHPKAKRM 8
DB 54 KHPKADRL 61
RESULT 9
Y132_METUA STANDARD; PRT; 220 AA.
AC Q57596;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein M0132.
GN M0132.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kervatage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geophagen N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii";
RL Science 273:1058-1073(1996).
CC -!- SIMILARITY: TO M.JANNA SCHII M01320 AND M01324.
CC -!- SIMILARITY: WITH TYPE I RESTRICTION SYSTEM ADENINE METHYLASES
(M SUBUNIT).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; U67470; AAB98113.1; -.
CC PIR; D64316; D64316.
CC TIGR; M0132; -.
CC InterPro; IPR003356; N6_DNA_Mtase.
DR Pfam; PF0384; N6_Mtase; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 220 AA; 25766 MW; 710DDAB4C7A47954 CRC64;
QY 1 KHPKAKRM 8
DB 54 KHPKADRL 61

Query Match 66.7%; Score 34; DB 1; Length 220;
Best Local Similarity 62.5%; Pred. No. 25;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 KHPKAKRM 8
DB 123 KHPKADRL 130
RESULT 10
Y132_METUA STANDARD; PRT; 433 AA.
AC Q57596;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein Y132.
GN Y132 OR IPA-93D OR BSU37600.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=168;
RX MEDLINE=98015417; PubMed=9353933;
RA Prescan E., Moszer I., Boursier L., Cruz Ramos H.C., De La Fuente V.,
RA Hulio M.-F., Lelong C., Schleich S., Sekowska A., Song B.H.,
RA Villani G., Kunst F., Danchin A., Glaser P.;
RA "The Bacillus subtilis genome from gerBC (311 degrees) to licR (334
degrees)";
RL Microbiology 143:3313-3328(1997).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertsch M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Ertz K.D., Errington J., Fabret C., Ferrazi E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Guim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppe G., Guy B.J., Haga K., Halech K., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klier-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Meliado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serron P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambut R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RA "The complete genome sequence of the Gram-positive bacterium Bacillus
subtilis";
RL Nature 390:249-256(1997).
[3]
RN SEQUENCE OF 1-97 FROM N.A.
RP STRAIN=168;
RX MEDLINE=95020537; PubMed=7934828;
RA Glaser P., Kunst F., Arnaud M., Coudart M.P., Gonzales W.,
RA Hulio M.-F., Ionescu M., Lubochinsky B., Marcelino L., Moszer I.,
RA Prescan E., Santana M., Schneider E., Schweizer J., Vertes A.,

RA Rapoport G., Danchin A.;
RT "Bacillus subtilis genome project: cloning and sequencing of the 97
kb region from 325 degrees to 333 degrees.";
RL Mol. Microbiol. 10:371-384(1993).
CC -!- SIMILARITY: TO M.GENITALIUM MG461.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; Z80355; CAB02494.1; -;
DR EMBL; Z99123; CAB15787.1; -;
DR EMBL; X73124; CAB15649.1; -;
DR PIR; G70056; G70056.
DR Subtilisin; BG10639; ywFO.
DR InterPro; IPR006674; HD.
DR InterPro; IPR003607; Met_phosphohydro.
DR Pfam; PF01966; HD; 1.
DR SMART; SM00471; HDC; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 433 AA; 50974 MW; D405C594102E4066 CRC64;

Query Match 66.7%; Score 34; DB 1; Length 433;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHPEAK 6
Db 418 KHPEAK 423
|||||

RESULT 11
YC20_YC20 META STANDARD; PRT; 578 AA.
AC Q58617;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein WJ1220.
GN WJ1220.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=JAI-1 / DSM 2661 / ATCC 43067;
MEDLINE=96337999; PubMed=6680807;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geochagan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.P., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RL Science 273:1058-1073(1996).
CC -!- SIMILARITY: TO M.JANNASCHII MJ0132 AND MJ0142.
CC -!- SIMILARITY: WITH TYPE I RESTRICTION SYSTEM ADENINE METHYLASES
CC (M SUBUNIT).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

CC or send an email to license@isb-sib.ch.
CC -----
CC EMBL; U67563; AB999225.1; -;
DR PIR; C64452; C64452.
DR TIGR; MJ1220; -;
DR InterPro; IPR003665; Methylase M.
DR InterPro; IPR002296; N12N6 mtfase.
DR InterPro; IPR003356; N6 DNA Mtase.
DR InterPro; IPR002052; N6 Mtase.
DR Pfam; PF02506; Methylase_M; 1.
DR Pfam; PF02384; N6 Mtase; 1.
DR PRINTS; PR00507; N12N6MTFRASE.
DR PROSITE; PS00092; N6 MTASE; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 578 AA; 66636 MW; 5339ED873EF8E9E2 CRC64;

Query Match 66.7%; Score 34; DB 1; Length 578;
Best Local Similarity 62.5%; Pred. No. 67;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KHPEAKRM 8
Db 481 KHPEVKKL 488
|||||

RESULT 12
T4BA_BACCO STANDARD; PRT; 637 AA.
ID T4BA_BACCO
AC Q07605;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Restriction enzyme BcglI alpha subunit [EC 3.1.21.-] [Includes:
DE Adenine-specific methyltransferase activity (EC 2.1.1.72)].
GN BCGIA.
OS Bacillus coagulans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1398;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 55055;
MEDLINE=93197166; PubMed=8451198;
RA Kong H., Morgan R.D., Maunus R.E., Schildkraut I.;
RT "A unique restriction endonuclease, BcglI, from Bacillus coagulans.";
RL Nucleic Acids Res. 21:987-991(1993).
RN [2]
RP SEQUENCE OF 1-14, AND CHARACTERIZATION.
RX STRAIN=ATCC 55055;
MEDLINE=94103292; PubMed=8276869;
RA Kong H., Roemer S.E., Waite-Rees P.A., Benner J.S., Wilson G.G.,
RA Nwankwo D.O.;
RT "Characterization of BcglI, a new kind of restriction-modification
system.";
RL J. Biol. Chem. 269:683-690(1994).
CC -!- FUNCTION: RECOGNIZES THE DOUBLE STRANDED SEQUENCE 5'-CGAN(6)TGC-3'
CC AND CLEAVES BILATERALLY AND SYMMETRICALLY OUTSIDE THE SEQUENCE TO
CC RELEASE A 34-BASE PAIR FRAGMENT. METHYLATION OF THE RECOGNITION
CC SEQUENCE OCCURS ON THE ADENINE IN EITHER ONE OR BOTH STRANDS.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA adenine = S-
CC adenosyl-L-homocysteine + DNA 6-methylaminopurine.
CC -!- COFACTOR: Magnesium.
CC -!- SUBUNIT: HETEROTRIMER OF TWO ALPHA AND ONE BETA SUBUNIT. BOTH
CC SUBUNITS ARE NECESSARY FOR DNA-BINDING.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L17341; AAA15626.1; -;

or send an email to license@isb-sib.ch.

CC
CC
CC
DR EMBL; AJ414143; CAC89354.1; -
DR EMBL; AR013971; AAM87226.1; -
DR PIR; AG0061; AG0061.
DR InterPro; IPR006145; Pseudou synth.
DR InterPro; IPR006224; Rlu synth.
DR Pfam; PF00849; Pseudou synth_2; 1.
DR PROSITE; PS01129; PSI_RLU; 1.
DR PROSITE; PS01129; PSI_RLU; 1.
KW tRNA processing; trna processing; Lyase; Complete proteome.
FT ACT SITE 51 51 BY SIMILARITY.
SQ SEQUENCE 206 AA; 23283 MW; F4951B5C5383F117 CRC64;

Query Match 64.7%; Score 33; DB 1; Length 206;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HPEAKRM 8
Db 170 HPEAKRM 176

RESULT 14

PROB_HAEIN STANDARD; PRT; 368 AA.

ID PROB_HAEIN STANDARD; PRT; 368 AA.

AC F43763;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glutamate 5-kinase (EC 2.7.2.11) (Gamma-glutamyl kinase) (GK).
DE Glutamate 5-kinase (EC 2.7.2.11) (Gamma-glutamyl kinase) (GK).
GN PROB OR HI0900.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kervatage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.M.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Pine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith K.O.,
RA Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
RD.";
RL Science 269:496-512(1995).
CC -!- FUNCTION: Catalyzes the transfer of a phosphate group to glutamate
CC to form glutamate 5-phosphate which rapidly cyclizes to 5-
CC oxoprolinate.
CC -!- CATALYTIC ACTIVITY: ATP + L-glutamate = ADP + L-glutamate 5-
CC phosphate.
CC -!- PATHWAY: Proline biosynthesis, first step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the glutamate 5-kinase family.
CC -!- SIMILARITY: Contains 1 PUA domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U32772; AAC22560.1; -
DR PIR; D64101; D64101.
DR TIGR; HI0900; -

DR HAMAP; MF 00456; -, 1.
DR InterPro; IPR001048; Aa_kinase.
DR InterPro; IPR001057; Glu_5kinase.
DR InterPro; IPR005715; ProB.
DR InterPro; IPR002478; PUA.
DR Pfam; PF00696; aakinasae; 1.
DR Pfam; PF01472; PUA; 1.
DR PRINTS; PRO0474; GLUSKINASE.
DR SMART; SM00359; PUA; 1.
DR TIGRfam; TIGR01027; ProB; 1.
DR PROSITE; PS00902; GLUTAMATE_5_KINASE; 1.
DR PROSITE; PS00890; PUA; 1.
KW Transferase; Kinase; Proline biosynthesis; Complete proteome.
FT DOMAIN 275 353
SQ SEQUENCE 368 AA; 40134 MW; E32B684A23709831 CRC64;

Query Match 64.7%; Score 33; DB 1; Length 368;
Best Local Similarity 66.7%; Pred. No. 66;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KHEPKRMP 9
|:||||:|
Db 180 KNPEAKLIP 188

RESULT 15
CP8B_RABIT
ID CP8B_RABIT STANDARD; PRT; 499 AA.
AC 002766;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 8B1 (EC 1.14.-.-) (CYPVIII1B1) (Sterol 12-alpha-hydroxylase) (7 alpha-hydroxy-4-cholesten-3-one 12-alpha-hydroxylase).
DE CYP8B1 OR CYP12.
GN Oryctolagus cuniculus (Rabbit).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=New Zealand white; TISSUE=Liver;
RX MEDLINE=97113033; PubMed=8943286;
RA Eggertsen G., Olin M., Andersson U., Ishida H., Kubota S., Hellman U.,
RA Okuda K.-I., Bjoekhem I.;
RT "Molecular cloning and expression of rabbit sterol 12alpha-hydroxylase";
RT J. Biol. Chem. 271:32269-32275 (1996).
RN [2]
RP SEQUENCE OF 1-15.
RC TISSUE=Liver;
RX MEDLINE=93016066; PubMed=1400444;
RA Ishida H., Noshiro M., Okuda K., Coon M.J.;
RT "Purification and characterization of 7 alpha-hydroxy-4-cholesten-3-one 12 alpha-hydroxylase";
RT J. Biol. Chem. 267:21319-21323 (1992).
CC -!- FUNCTION: Involved in bile acid synthesis and is responsible for the conversion of 7 alpha-hydroxy-4-cholesten-3-one into 7 alpha, 12 alpha-dihydroxy-4-cholesten-3-one. Responsible for the balance between formation of cholic acid and chenodeoxycholic acid. Has a rather broad substrate specificity including a number of 7-alpha-hydroxylated C27 steroids.
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum membrane. May be anchored to the membrane via a single transmembrane domain.
CC -!- TISSUE SPECIFICITY: Liver.
CC -!- INDUCTION: By starvation.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y08269; CAA69594.1; -.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 2.
DR PROSITE; PS00086; CYTOCHROME_P450; FALSE NEG.
KW Oxidoreductase; Monooxygenase; Electron transport; Transmembrane;
KW Heme; Microsome; Endoplasmic reticulum.
FT INIT MET 0
FT TRANSMEM 1 20 POTENTIAL.
FT METAL 438 438 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 499 AA; 57363 MW; 8FE56CADEDED3C3A CRC64;

Query Match 64.7%; Score 33; DB 1; Length 499;
Best Local Similarity 85.7%; Pred. No. 90;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KHEPKR 7
|:||||:|
Db 298 KHEPKMR 304

Search completed: April 19, 2004, 11:52:55
Job time : 1.65651 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 0.6561 seconds

(without alignments)
713.823 Million cell updates/sec

Title: US-09-832-929-18_COPY_478_486

Perfect score: 45

Sequence: 1 TESLVNRRP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	600	1	ALBU_MACMU
2	45	100.0	605	1	ALBU_PIG
3	45	100.0	607	1	ALBU_BOVIN
4	45	100.0	607	1	ALBU_SHEEP
5	45	100.0	608	1	ALBU_FELCA
6	45	100.0	609	1	ALBU_HUMAN
7	41	91.1	608	1	ALBU_CANFA
8	36	80.0	1241	1	ATB4_HUMAN
9	35	77.8	608	1	ALBU_RABIT
10	33	73.3	1112	1	ATB2_OREMO
11	33	73.3	1203	1	ATB4_RAT
12	32	71.1	73	1	LGEM_BOVIN
13	32	71.1	315	1	FLIM_BUCAL
14	32	71.1	316	1	MRAW_STRUM
15	32	71.1	607	1	ALBU_HORSE
16	31	68.9	317	1	GSHB_PSEPK
17	31	68.9	422	1	FEW2_HUMAN
18	31	68.9	422	1	FEW2_MOUSE
19	31	68.9	525	1	ALGG_AZOVI
20	31	68.9	609	1	FETA_GORGO
21	31	68.9	609	1	FETA_HORSE
22	31	68.9	609	1	FETA_HUMAN
23	31	68.9	609	1	FLK2_PANTR
24	31	68.9	877	1	CLK2_CABEL
25	30	66.7	240	1	TPIS_GLOVI
26	30	66.7	465	1	DGT2_LISIN
27	30	66.7	465	1	DGT2_LISMO
28	30	66.7	477	1	LE21_BRAJA
29	30	66.7	478	1	ARCDC_CLOPE
30	30	66.7	502	1	C72K_ARATH
31	30	66.7	614	1	DNAK_ODOSI
32	30	66.7	704	1	VPS1_YEAST
33	30	66.7	855	1	XAB2_HUMAN

34	30	66.7	855	1	XAB2_MOUSE
35	30	66.7	855	1	XAB2_RAT
36	30	66.7	882	1	CTIB_FUSO
37	30	66.7	993	1	YAJ1_SCHPO
38	30	66.7	1030	1	PEX6_YEAST
39	30	66.7	1220	1	ATB1_PIG
40	30	66.7	1220	1	ATB3_HUMAN
41	30	66.7	1225	1	SMC1_YEAST
42	30	66.7	1258	1	ATB1_HUMAN
43	30	66.7	1258	1	ATB1_RAT
44	30	66.7	1258	1	ATB3_RAT
45	29	64.4	147	1	PHIT_HUMAN

ALIGNMENTS

RESULT 1
ALBU_MACMU
ID ALBU_MACMU STANDARD; PRT; 600 AA.
AC Q28522;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor (Fragment).
GN ALB.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=93211971; PubMed=8460152;
RX Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,
RA Deulet J., Putnam F.W.;
RT "cDNA and protein sequence of polymorphic macaque albumins that differ
in bilirubin binding";
RL Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Plasma.
CC -1- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -1- SIMILARITY: Contains 3 albumin domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).

EMBL: M90463; AAA36906.1; -
PIR: A47391; A47391.
HSP: P02768; I57B.
InterPro: IPR000284; Serum_albumin.
Pfam: PF00273; transport_prot; 3.
PRINTS: PR00802; SERUMALBUMIN.
ProDom: PD002486; Serum_albumin; 1.
SMART: SM00103; ALBUMIN; 3.
PROSITE: PS00212; ALBUMIN; 3.
Metal-binding; Lipid-binding; Repeat; Signal; Copper.
NON_TER 1 1
SIGNAL <1 10 BY SIMILARITY.
PROPEP 11 16 BY SIMILARITY.
FT CHAIN 17 600 SERUM ALBUMIN.
FT DOMAIN 17 197 ALBUMIN 1.
FT DOMAIN 204 389 ALBUMIN 2.
FT DOMAIN 396 587 ALBUMIN 3.

O9dcd2	mus musculus
O99pk0	rattus norv
P52959	fusarium so
O09901	schizosacch
P33760	saccharomyc
P23220	sus scrofa
Q16720	homo sapien
P32908	saccharomyc
P20020	homo sapien
P11505	rattus norv
O64588	rattus norv
P49789	homo sapien

```
FT METAL 19 19
FT BINDING 256 256
FT DISULFID 69 78
FT DISULFID 91 107
FT DISULFID 106 117
FT DISULFID 140 185
FT DISULFID 184 193
FT DISULFID 216 262
FT DISULFID 261 269
FT DISULFID 281 295
FT DISULFID 294 305
FT DISULFID 332 377
FT DISULFID 376 385
FT DISULFID 408 454
FT DISULFID 453 464
FT DISULFID 477 493
FT DISULFID 492 503
FT DISULFID 530 575
FT DISULFID 574 583
SQ SEQUENCE 600 AA; 67890 MW; E45C871A670E740B CRC64;

Query Match 100.0%; Score 45; DB 1; Length 600;
Best Local Similarity 100.0%; Pred. No. 0.085;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TESLVNRRP 9
Db 494 TESLVNRRP 502

RESULT 2
ALBU_PIG STANDARD; PRT; 605 AA.
ID ALBU_PIG
AC P08835; Q29018;
DT 01-NOV-1998 (Rel. 09, Created)
DT 01-NOV-1998 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor (Fragment).
GN ALB
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RC SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89016582; PubMed=3174440;
RA Baldwin G.S., Weinstock J.;
RT "Nucleotide sequence of porcine liver albumin.";
RL Nucleic Acids Res. 16:9045-9045(1988).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X12422; CAA30970.1; -
CC EMBL; X36787; AAA30988.1; -
CC FIR; S01382; ABPGS.
CC HSP; P02768; 1E7H.
CC InterPro; IPR000264; Serum_albumin.
CC Pfam; PF00273; transport_prot; 3.
```

```
DR PRINTS; P00802; SERUMALBUMIN.
DR ProDom; P002486; Serum_albumin; 1.
DR SMART; SMO0103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT NON_TER 1 1
FT SIGNAL <1 16
FT PROPEP 17 22
FT CHAIN 23 605
FT DOMAIN 23 202
FT DOMAIN 209 394
FT DOMAIN 401 592
FT METAL 31 31
FT DISULFID 75 84
FT DISULFID 97 113
FT DISULFID 112 123
FT DISULFID 145 190
FT DISULFID 189 198
FT DISULFID 221 267
FT DISULFID 266 274
FT DISULFID 286 300
FT DISULFID 299 310
FT DISULFID 337 382
FT DISULFID 381 390
FT DISULFID 413 459
FT DISULFID 458 469
FT DISULFID 482 498
FT DISULFID 497 508
FT DISULFID 535 580
FT DISULFID 579 588
FT CONFLICT 562 562
SQ SEQUENCE 605 AA; 69410 MW; 3E55B0DD1A1F4FF CRC64;

Query Match 100.0%; Score 45; DB 1; Length 605;
Best Local Similarity 100.0%; Pred. No. 0.085;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TESLVNRRP 9
Db 499 TESLVNRRP 507

RESULT 3
ALBU_BOVIN STANDARD; PRT; 607 AA.
ID ALBU_BOVIN
AC P02769; O02787;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Bos d 6).
GN ALB
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RC SEQUENCE FROM N.A.
RA Holowachuk E.W., Stoltenberg J.K., Read R.G., Peters T. Jr.;
RL Submitted (AUG-1991) to the EMBL/GenBank/DBJ databases.
RN [2]
RC SEQUENCE FROM N.A., AND VARIANT THR-214.
RC TISSUE=Liver;
RA Barry T., Power S., Gannon F.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RC SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Hilger C., Grigioni F., de Beaufort C., Michel G., Hentges F.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RC SEQUENCE FROM N.A., AND VARIANT THR-214.
RA Wu H.T., Huang M.C.;
```

RT "The complete cDNA sequence of bovine serum albumin.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-32.
RX MEDLINE=80024278; PubMed=488109;
RA McGilivray R.T.A., Chung D.W., Davie E.W.;
RT "Biosynthesis of bovine plasma proteins in a cell-free system. Amino-
terminal sequence of prealbumin.";
RL Eur. J. Biochem. 98:477-485(1979).
RN [6]
RP SEQUENCE OF 25-424 AND 429-607, AND VARIANT THR-214.
RA Brown J.R.;
RT "Structure of bovine serum albumin.";
RL Fed. Proc. 34:591-591(1975).
RN [7]
RP REVISIONS TO 190-195.
RA Brown J.R.;
RT Submitted (APR-1975) to the PIR data bank.
RN [8]
RP SEQUENCE OF 402-433.
RX MEDLINE=8203364; PubMed=7283978;
RA Reed R.G., Putnam F.W., Peters T. Jr.;
RT "Sequence of residues 400-403 of bovine serum albumin.";
RL Biochem. J. 191:867-868(1980).
RN [9]
RP SEQUENCE OF 19-28.
RX MEDLINE=77134075; PubMed=843354;
RA Patterson J.E., Geller D.M.;
RT "Bovine microsomal albumin: amino terminal sequence of bovine
proalbumin.";
RL Biochem. Biophys. Res. Commun. 74:1220-1226(1977).
RN [10]
RP SEQUENCE, AND REVISIONS TO 118-119 AND 180.
RX MEDLINE=91083649; PubMed=2260975;
RA Hirayama K., Akashi S., Furiya M., Fukuhara K.-I.;
RT "Rapid confirmation and revision of the primary structure of bovine
serum albumin by ESIMS and Frit-FAB LC/MS.";
RL Biochem. Biophys. Res. Commun. 173:639-646(1990).
RN [11]
RP SEQUENCE OF 25-41.
RX MEDLINE=88267456; PubMed=3389500;
RA Hsieh J.C., Lin F.P., Tam M.F.;
RT "Electroblotting onto glass-fiber filter from an analytical
isoelectrofocusing gel: a preparative method for isolating proteins
for N-terminal microsequencing.";
RL Anal. Biochem. 170:11-8(1988).
RN [12]
RP SEQUENCE OF 437-451.
RX Vilbois F.;
RA Submitted (AUG-1998) to Swiss-Prot.
RN [13]
RP DISULFIDE BONDS.
RA Brown J.R.;
RT "Structure of serum albumin: disulfide bridges.";
RL Fed. Proc. 33:1389-1389(1974).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; M73993; AAA51411.1; --
DR EMBL; X58989; CAA41735.1; --
DR EMBL; Y17769; CAA76847.1; --
DR EMBL; AF542068; AAN17824.1; --
DR HSSP; P02768; 1E7B
DR InterPro; IPR00264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR PRODOM; PD002486; Serum_albumin; 1.
DR SMART; SMO0103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen;
KW Polymorphism.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 607 SERUM ALBUMIN.
FT DOMAIN 25 204 ALBUMIN 1.
FT DOMAIN 211 396 ALBUMIN 2.
FT DOMAIN 403 594 ALBUMIN 3.
FT METAL 27 27 COPPER (BY SIMILARITY).
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 147 192
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 288 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461
FT DISULFID 460 471
FT DISULFID 484 500
FT DISULFID 499 510
FT DISULFID 537 582
FT DISULFID 581 590
FT VARIANT 214 214
FT CONFLICT 302 302
FT CONFLICT 304 305
FT CONFLICT 324 324
FT CONFLICT 394 395
FT CONFLICT 437 437
FT CONFLICT 493 494
FT CONFLICT 607 607
SQ SEQUENCE 607 AA; 69293 MW; 39167DFE768585D4 CRC64;
A -> T.
C -> K (IN REF. 6).
KP -> PC (IN REF. 6).
N -> D (IN REF. 6).
ST -> TS (IN REF. 6).
K -> R (IN REF. 12).
SE -> ES (IN REF. 6).
Query Match 100.0%; Score 45; DB 1; Length 607;
Best Local Similarity 100.0%; Pred.No. 0.086;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TESLVNRRP 9
DB 501 TESLVNRRP 509
RESULT 4
ALBU_SHEEP
ID ALBU_SHEEP STANDARD; PRT; 607 AA.
AC P14639;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=9009888; PubMed=2602160;

RA Brown W.M., Dziegielewska K.M., Foreman R.C., Saunders N.R.;
 RT "Nucleotide and deduced amino acid sequence of sheep serum albumin.";
 RL Nucleic Acids Res. 17:10495-10495(1989).
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 CC hormones, bilirubin and drugs. Its main function is the regulation
 CC of the colloidal osmotic pressure of blood.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Plasma.
 CC -1- SIMILARITY: Belongs to the ALB/APP/VDB family.
 CC -1- SIMILARITY: Contains 3 albumin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X17055; CAA34903.1; -;
 CC PIR; S06936; ABSHS.
 CC HSP; P02768; 1E7B.
 CC InterPro: IPR000264; Serum_albumin.
 CC Pfam: PF00273; transport_prot; 3.
 CC PRINTS; PR00802; SERUMALBUMIN.
 CC ProDom; PD002486; Serum_albumin; 1.
 CC SMART; SM00103; ALBUMIN; 3.
 CC PROSITE; PS00212; ALBUMIN; 3.
 CC Metal-binding; Lipid-binding; Repeat; Signal; Copper.
 FT SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 607
 FT DOMAIN 25 204
 FT DOMAIN 211 396
 FT DOMAIN 403 594
 FT METAL 27 27
 FT METAL 77 86
 FT DISULFID 77 86
 FT DISULFID 99 115
 FT DISULFID 114 125
 FT DISULFID 147 192
 FT DISULFID 191 200
 FT DISULFID 223 269
 FT DISULFID 268 276
 FT DISULFID 288 302
 FT DISULFID 301 312
 FT DISULFID 339 384
 FT DISULFID 383 392
 FT DISULFID 415 461
 FT DISULFID 460 471
 FT DISULFID 484 500
 FT DISULFID 499 510
 FT DISULFID 537 582
 FT DISULFID 581 590
 SQ SEQUENCE 607 AA; 69188 MW; 84979A87F9B86596 CRC64;
 Query Match 100.0%; Score 45; DB 1; Length 607;
 Best Local Similarity 100.0%; Pred. No. 0.086;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
 DB 501 TESLVNRRP 509

RESULT 5
 ALBU_FELCA STANDARD; PRT; 608 AA.
 ID ALBU_FELCA
 AC P49064;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Serum albumin precursor (Allergen Fel d 2).

GN ALB.
 OS Felis silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
 OX NCBI_TaxID=9685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96194824; PubMed=8647469;
 RA Hilger C., Grigioni F., Kohnen M., Hentges F.;
 RT "Sequence of the gene encoding cat (Felis domesticus) serum albumin.";
 RL Gene 169:295-296(1996).
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 CC hormones, bilirubin and drugs. Its main function is the regulation
 CC of the colloidal osmotic pressure of blood.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Plasma.
 CC -1- ALLERGEN: Causes an allergic reaction in human.
 CC -1- SIMILARITY: Belongs to the ALB/APP/VDB family.
 CC -1- SIMILARITY: Contains 3 albumin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X84842; CAA59279.1; -;
 CC PIR; JC4660; S57632.
 CC HSP; P02768; 1E7B.
 CC InterPro: IPR000264; Serum_albumin.
 CC Pfam: PF00273; transport_prot; 3.
 CC PRINTS; PR00802; SERUMALBUMIN.
 CC ProDom; PD002486; Serum_albumin; 1.
 CC SMART; SM00103; ALBUMIN; 3.
 CC PROSITE; PS00212; ALBUMIN; 3.
 CC Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
 FT SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 608
 FT DOMAIN 25 205
 FT DOMAIN 212 397
 FT DOMAIN 404 595
 FT METAL 27 27
 FT METAL 77 86
 FT DISULFID 77 86
 FT DISULFID 99 115
 FT DISULFID 114 125
 FT DISULFID 148 193
 FT DISULFID 192 201
 FT DISULFID 224 270
 FT DISULFID 269 277
 FT DISULFID 289 303
 FT DISULFID 302 313
 FT DISULFID 340 385
 FT DISULFID 384 393
 FT DISULFID 416 462
 FT DISULFID 461 472
 FT DISULFID 485 501
 FT DISULFID 500 511
 FT DISULFID 538 583
 FT DISULFID 582 591
 SQ SEQUENCE 608 AA; 68659 MW; 07E629CAC5F60E5F CRC64;
 Query Match 100.0%; Score 45; DB 1; Length 608;
 Best Local Similarity 100.0%; Pred. No. 0.086;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
 DB 502 TESLVNRRP 510

RESULT 6
ALBU HUMAN STANDARD: PRT: 609 AA.
ID ALBU HUMAN STANDARD: PRT: 609 AA.
AC P02768; O95574; Q13140; Q9P157; Q9P117; Q9UJ20;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=86196112; PubMed=3009475;
RA Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,
RA Beattis W.G., Dugaiczky A.,
RT "Molecular structure of the human albumin gene is revealed by
RT nucleotide sequence within q11-22 of chromosome 4.";
RL J. Biol. Chem. 261:6747-6757(1986).
RN [2]
RN SEQUENCE FROM N.A., AND VARIANT LYS-420.
RX MEDLINE=82081882; PubMed=6171778;
RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,
RA Najarian R.C., Seeburg P.H., Wion K.L.;
RT "The sequence of human serum albumin cDNA and its expression in E.
RT coli.";
RL Nucleic Acids Res. 9:6103-6114(1981).
RN [3]
RN SEQUENCE FROM N.A.
RX MEDLINE=82105994; PubMed=6275391;
RA Dugaiczky A., Law S.W., Dennison O.E.;
RT "Nucleotide sequence and the encoded amino acids of human serum
RT albumin mRNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).
RN [4]
RN SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RN SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).
RC TISSUE=Fetal liver;
RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,
RA Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;
RT "Functional prediction of the coding sequences of 121 new genes
RT deduced by analysis of cDNA clones from human fetal liver.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [6]
RN SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.
RA Huang M.C., Wu H.T.;
RT "The cDNA sequences of human serum albumin.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [7]
RN SEQUENCE FROM N.A.
RC TISSUE=Liver, and Skeletal muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenon C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S., Lequellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Heiton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [8]
RN SEQUENCE OF 25-609.
RX MEDLINE=76187907; PubMed=1225573;
RA Meloun B., Moravsek L., Kostka V.;
RT "Complete amino acid sequence of human serum albumin.";
RL FEBS Lett. 58:134-137(1975).
RN [9]
RN SEQUENCE OF 25-609.
RA Brown J.R., Shockley P., Behrens P.Q.;
RL (in) Bing D.H. (eds.);
RT The chemistry and physiology of the human plasma proteins, pp.23-40,
RL Pergamon Press, New York (1979).
RN [10]
RN SEQUENCE OF 1-455 FROM N.A.
RC TISSUE=Liver;
RA Menaya J., Parrilla R., Ayuso M.S.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [11]
RN SEQUENCE OF 1-26 FROM N.A.
RX MEDLINE=86140099; PubMed=2419329;
RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;
RT "The human albumin gene. Characterization of the 5' and 3' flanking
RT regions and the polymorphic gene transcripts.";
RL J. Biol. Chem. 261:3244-3251(1986).
RN [12]
RN SEQUENCE OF 222-229.
RX MEDLINE=76257808; PubMed=955075;
RA Walker J.E.;
RT "Lysine residue 199 of human serum albumin is modified by
RT acetylsalicylic acid.";
RL FEBS Lett. 66:173-175(1976).
RN [13]
RN SEQUENCE OF 25-44 AND 480-499.
RC TISSUE=Heart;
RX MEDLINE=95203287; PubMed=7895732;
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
RT "The human myocardial two-dimensional gel protein database: update
RT 1994.";
RL Electrophoresis 15:1459-1465(1994).
RN [14]
RN DISULFIDE BONDS.
RA Sabar M.A., Stockbauer P., Moravsek L., Meloun B.;
RT "Disulfide bonds in human serum albumin.";
RL Collect. Czech. Chem. Commun. 42:564-579(1977).
RN [15]
RN BILIRUBIN-BINDING SITE.
RX MEDLINE=78186630; PubMed=656055;
RA Jacobsen C.;
RT "Lysine residue 240 of human serum albumin is involved in high-
RT affinity binding of bilirubin.";
RL Biochem. J. 171:453-459(1978).
RN [16]
RN VARIANT CANTERBURY ASN-337.
RX MEDLINE=87157744; PubMed=3828358;
RA Brennan S.O., Herbert P.;
RT "Albumin Canterbury (313 Lys->Asn). A point mutation in the second
RT domain of serum albumin.";
RL Biochim. Biophys. Acta 912:191-197(1987).
RN [17]
RN VARIANTS NAG-2 AND NAG-3.
RX MEDLINE=88068523; PubMed=3479777;
RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,
RA Satoh C., Neel J.V.;
RT "Amino acid substitutions in inherited albumin variants from
RT Amerindian and Japanese populations.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).
RN [18]

RP VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.
RX MEDLINE=89345611; PubMed=2762316;
RA Arai K., Madison J., Huse K., Ishioka N., Satoh C., Fujita M.,
RA Neel J.V., Sakurabayashi I., Putnam F.W.;
RT "Point substitutions in Japanese allolbumins.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).
RN [19]
RP VARIANTS MANAUS; OSAKA; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.
RX MEDLINE=90115905; PubMed=2404284;
RA Arai K., Madison J., Shimizu A., Putnam F.W.;
RT "Point substitutions in albumin genetic variants from Asia.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).
RN [20]
RP DESCRIPTION OF VARIANT REDHILL.
RX MEDLINE=90115852; PubMed=2104980;
RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;
RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of
human serum albumin whose precursor has an aberrant signal peptidase
cleavage site.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).
RN [21]
RP VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.
RX MEDLINE=91062352; PubMed=2247440;
RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,
RA Watkins S., Putnam F.W.;
RT "Mutations in genetic variants of human serum albumin found in
Italy.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).
RN [22]
RP VARIANT VENEZIA.
RX MEDLINE=91296740; PubMed=2068071;
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,
RA Minchiotti L., Putnam F.W.;
RT "A donor splice mutation and a single-base deletion produce two
carboxyl-terminal variants of human serum albumin.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).
RN [23]
RP VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;
RX MEDLINE=92052189; PubMed=1945412;
RA Madison J., Arai K., Reid R.D., Kyle R.A., Watkins S., Davis E.,
RA Matsuda Y.-I., Amaki I., Putnam F.W.;
RT "Genetic variants of serum albumin in Americans and Japanese.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).
RN [24]
RP VARIANT CASEBROOK ASN-518.
RX MEDLINE=91316157; PubMed=1859851;
RA Peach R.J., Brennan S.O.;
RT "Structural characterization of a glycoprotein variant of human serum
albumin: albumin Casebrook (494 Asp-->Asn).";
RL Biochim. Biophys. Acta 1097:49-54(1991).
RN [25]
RP VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.
RX MEDLINE=92190239; PubMed=1347703;
RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,
RA Rochu D., Porta F.;
RT "Two allolbumins with identical electrophoretic mobility are produced
Query Match 100.0%; Score 45; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.086;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
Db 502 TESLVNRRP 510

RESULT 7
ID ALBU-CANFA STANDARD; PRT; 608 AA.
AC P49822; 077705; 09TS24;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Can f 3).
GN ALB.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_taxid=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Beagle; TISSUE=Liver;
RA Hilger C.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Liver;
RA MEDLINE=20148667; PubMed=10698848;
RA Pandjaitan B., Swoboda I., Brandefsky-Fichler F., Rumpold H.,
RA Valenta R., Spitzauer S.;
RT "Escherichia coli expression and purification of recombinant dog
albumin, a cross-reactive animal allergen.";
RL J. Allergy Clin. Immunol. 105:279-285(2000).
RN [3]
RP SEQUENCE OF 25-48.
RX MEDLINE=75011422; PubMed=4414013;
RA Dixon J.W., Sarkar B.;
RT "Isolation, amino acid sequence and copper(II)-binding properties of
peptide (1-24) of dog serum albumin.";
RL J. Biol. Chem. 249:5872-5877(1974).
RN [4]
RP SEQUENCE OF 25-38.
RX TISSUE=Heart;
RA MEDLINE=98163340; PubMed=9504812;
RA Dunn M.J., Corbett J.M., Wheeler C.H.;
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
dog heart proteins.";
RL Electrophoresis 18:2795-2802(1997).
RN [5]
RP SEQUENCE OF 215-478 FROM N.A.
RX TISSUE=Salivary gland;
RA MEDLINE=94201492; PubMed=7512102;
RA Spitzauer S., Schweizer C., Sperr W.R., Pandjaitan B., Valent P.,
RA Muehl S., Ebner C., Scheiner O., Kraft D., Rumpold H.;
RT "Molecular characterization of dog albumin as a cross-reactive
allergen.";
RL J. Allergy Clin. Immunol. 93:614-627(1994).
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Plasma.
CC -1- ALLERGEN: Causes an allergic reaction in human.
CC -1- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC -1- SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; AJ133489; CAB64567.1; -;
CC EMBL; Y1737; CAB75841.1; -;
CC EMBL; S72946; AAB30434.1; -;
CC HSSP; P02768; 1E7B.
CC HSC-2DPAGE; P49822; DOG.
CC InterPro; IPR000264; Serum_albumin.
CC Pfam; PF00273; transport_prot; 3.
CC PRINTS; PF00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum albumin; 1.
CC SMART; SM00103; ALBUMIN; 3.

DR PROSITE; PS00212; ALBUMIN; 3.
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
 FT SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 608
 FT DOMAIN 25 205
 FT DOMAIN 212 397
 FT DOMAIN 404 595
 FT METAL 27 27
 FT DISULFID 77 86
 FT DISULFID 99 115
 FT DISULFID 114 125
 FT DISULFID 148 193
 FT DISULFID 192 201
 FT DISULFID 224 270
 FT DISULFID 269 277
 FT DISULFID 289 303
 FT DISULFID 302 313
 FT DISULFID 340 395
 FT DISULFID 384 393
 FT DISULFID 416 462
 FT DISULFID 461 472
 FT DISULFID 485 501
 FT DISULFID 500 511
 FT DISULFID 538 583
 FT DISULFID 582 591
 FT CONFLICT 1 26
 FT CONFLICT 146 146
 FT CONFLICT 206 206
 FT CONFLICT 349 349
 FT CONFLICT 359 359
 FT CONFLICT 448 448
 FT CONFLICT 474 474
 FT CONFLICT 474 474
 SQ SEQUENCE 608 AA; 68606 MW; 3CF1C8FF7DBDFC06 CRC64;
 Query Match 91.1%; Score 41; DB 1; Length 608;
 Best Local Similarity 88.9%; Pred. No. 0.62;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TESLVNRRP 9
 Db 502 SESLVNRRP 510
 RESULT 8
 ID ATB4_HUMAN STANDARD; PRT; 1241 AA.
 AC P23634; Q13450; Q13452; Q13455; Q16817;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Plasma membrane calcium-transporting ATPase 4 (EC 3.6.3.8) (PMCA4)
 DE (Plasma membrane calcium pump isoform 4) (Plasma membrane calcium
 DE ATPase isoform 4).
 GN ATP2B4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A. (ISOFORM XB), AND PARTIAL SEQUENCE.
 RN TISSUE=Erythrocyte;
 RX MEDLINE=90153913; PubMed=2137451;
 RA Strehler E.E., James P., Fischer R., Heim R., Vorherr T.E.,
 RA Piloteo A.G., Penniston J.T., Carafoli E.,
 RT "Peptide sequence analysis and molecular cloning reveal two calcium
 RT pump isoforms in the human erythrocyte membrane.";
 RL J. Biol. Chem. 265:2835-2842(1990).
 [2]
 RN SEQUENCE FROM N.A. (ISOFORM XA).
 RP TISSUE=Fetal brain;
 RC MEDLINE=92165787; PubMed=1531651;
 RX

RA Brandt P., Neve R.L., Kammesheidt A., Rhoads R.E., Vanaman T.C.;
 RT "Analysis of the tissue-specific distribution of mRNAs encoding the
 RT plasma membrane calcium-pumping ATPases and characterization of an
 RT alternately spliced form of PMCA4 at the cDNA and genomic levels.";
 RL J. Biol. Chem. 267:4376-4385(1992).
 [3]
 RN PARTIAL SEQUENCE FROM N.A. (ISOFORMS A; B; D AND K).
 RP TISSUE=Heart muscle;
 RX MEDLINE=96276517; PubMed=8700162;
 RA Santiago-Garcia J., Mas-Oliva J., Saavedra D., Zarain-Herzberg A.;
 RT "Analysis of mRNA expression and cloning of a novel plasma membrane
 RT Ca(2+)-ATPase splice variant in human heart.";
 RL Mol. Cell. Biochem. 155:173-182(1996).
 [4]
 RN ALTERNATIVE SPLICING (ISOFORMS X AND Z).
 RP TISSUE=Heart;
 RX MEDLINE=94064681; PubMed=8245032;
 RA Stauffer T.P., Hilfinger H., Carafoli E., Strehler E.E.;
 RT "Quantitative analysis of alternative splicing options of human plasma
 RT membrane calcium pump genes.";
 RL J. Biol. Chem. 268:25993-26003(1993).
 [5]
 RN ERATUM.
 RP MEDLINE=95081164; PubMed=7989379;
 RX Stauffer T.P., Hilfinger H., Carafoli E., Strehler E.E.;
 RA J. Biol. Chem. 269:32022-32022(1994).
 CC -1- FUNCTION: This magnesium-dependent enzyme catalyzes the hydrolysis
 CC of ATP coupled with the transport of calcium out of the cell.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + Ca(2+) (Cis) = ADP + phosphate +
 CC Ca(2+) (Trans).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Comment=There is a combination of two alternatively spliced
 CC domains at N-terminal site A (X and Z) and at C-terminal site
 CC B/C (A, B, D and K). The splice sites have mostly been studied
 CC independently. Full isoforms so far detected are isoform XA and
 CC isoform XB. Experimental confirmation may be lacking for some
 CC isoforms;
 CC Name=XD; Synonyms=AIIICV;
 CC IsoId=P23634-1; Sequence=Displayed;
 CC Name=XA; Synonyms=AIIICII;
 CC IsoId=P23634-2; Sequence=VSP_000405;
 CC Name=ZA; Synonyms=ALCII;
 CC IsoId=P23634-3; Sequence=VSP_000402, VSP_000405;
 CC Name=XK; Synonyms=XG;
 CC IsoId=P23634-4; Sequence=VSP_000403, VSP_000405;
 CC Name=ZK; Synonyms=ZG;
 CC IsoId=P23634-5; Sequence=VSP_000402, VSP_000403, VSP_000405;
 CC Name=XB; Synonyms=AIIICI;
 CC IsoId=P23634-6; Sequence=VSP_000404;
 CC Name=ZB; Synonyms=ALCI;
 CC IsoId=P23634-7; Sequence=VSP_000402, VSP_000404;
 CC Name=ZD; Synonyms=AIIICV;
 CC IsoId=P23634-8; Sequence=VSP_000402;
 CC -1- TISSUE SPECIFICITY: Isoform XB is the most abundant isoform and is
 CC expressed ubiquitously. Isoforms containing segment Z have only
 CC been detected in heart, while isoforms containing segment A have
 CC been found in heart, stomach and brain cortex.
 CC -1- SIMILARITY: Belongs to the cation transport ATPases family (P-type
 CC ATPases). Subfamily IIB.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M25874; AAA50819.1; -;
 DR EMBL; M83363; AAA36455.1; -;
 DR EMBL; U42026; AAB17577.1; -;

DR	EMBL; U42061; AAB17578.1; -.
DR	EMBL; U42062; AAB17579.1; -.
DR	EMBL; U42378; AAB17580.1; -.
DR	PIR; A35547; A35547.
DR	HSSP; P04191; 1EUL.
DR	Genew; HGNC:817; ATP2B4.
DR	MIM; 108732; -.
DR	GO; GO:0005887; C:integral to plasma membrane; TAS.
DR	GO; GO:0005388; F:calcium-transporting ATPase activity; TAS.
DR	GO; GO:0006810; P:transport; TAS.
DR	InterPro; IPRC06408; ATPase-IIB_Ca.
DR	InterPro; IPRC01757; ATPase-EI-E2.
DR	InterPro; IPRC006068; Cation_ATPase_C.
DR	InterPro; IPRC004014; Cation_ATPase_N.
DR	InterPro; IPRC008250; EI-E2_ATPase_reg.
DR	InterPro; IPRC005834; Hydrolase.
DR	Pfam; PF006689; Cation_ATPase_C; 1.
DR	Pfam; PF00690; Cation_ATPase_N; 1.
DR	Pfam; PF001122; EI-E2_ATPase; 1.
DR	Pfam; PF00702; Hydrolase; 1.
DR	PRINTS; PR001119; CATATPASE.
DR	TIGRFAMS; TIGR01517; ATPase-IIB_Ca; 1.
DR	TIGRFAMS; TIGR01494; ATPase-EI-E2; 1.
DR	PROSITE; PS00154; ATPase_EI-E2; 1.
KW	Hydrolase; Calcium transport; Transmembrane; Phosphorylation;
KW	Magnesium; Metal-binding; ATP-binding; Calmodulin-binding;
KW	Multigene family; Alternative splicing.
KW	DOMAIN 1 92 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 93 113 POTENTIAL.
FT	DOMAIN 114 150 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 151 171 POTENTIAL.
FT	DOMAIN 172 356 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 357 376 POTENTIAL.
FT	DOMAIN 377 409 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 410 427 POTENTIAL.
FT	DOMAIN 428 840 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 841 860 POTENTIAL.
FT	DOMAIN 861 870 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 871 891 POTENTIAL.
FT	DOMAIN 892 911 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 912 934 POTENTIAL.
FT	DOMAIN 935 952 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 953 974 POTENTIAL.
FT	DOMAIN 975 993 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 994 1015 POTENTIAL.
FT	DOMAIN 1016 1025 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 1026 1047 POTENTIAL.
FT	DOMAIN 1048 1241 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN 1241 1241 POLY-LYS.
FT	DOMAIN 1241 1241 CALMODULIN-BINDING SUBDOMAIN A (BY SIMILARITY).
FT	DOMAIN 1086 1103 CALMODULIN-BINDING SUBDOMAIN B (BY SIMILARITY).
FT	DOMAIN 1104 1113 POLY-GLU.
FT	DOMAIN 1186 1190 PHOSPHORYLATION (BY SIMILARITY).
FT	MOD RES 465 465 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
FT	MOD RES 1102 1102 MAGNESIUM (BY SIMILARITY).
FT	METAL 785 785 MAGNESIUM (BY SIMILARITY).
FT	METAL 789 789 Missing (in isoform ZA, isoform ZK, isoform ZB and isoform ZD).
FT	VARSPLIC 301 312 /FTID-VSP_000402.
FT	VARSPPLIC 1009 1044 Missing (in isoform XK and isoform ZK).
FT	VARSPPLIC 1104 1139 /FTID-VSP_000403.
FT	VARSPPLIC 1140 1241 Missing (in isoform XB and isoform ZB).
FT	VARSPPLIC 1140 1241 /FTID-VSP_000404.
FT	VARSPPLIC 1140 1241 IKVLLDERENPDKASKFGTGVLLDGDEVTPTANTNNNAVD
FT	VARSPPLIC 1140 1241 TPFLAEEENPDKASKFGTGVLLDGDEVTPTANTNNNAVD
FT	VARSPPLIC 1140 1241 CNOVOLPOSDLSLOSLETSV -> VAVAPVKSSPTTSVPAP
FT	VARSPPLIC 1140 1241 SSPPMNQSGQSVF (in isoform XA, isoform XK, isoform ZB and isoform ZK).
FT	VARSPPLIC 1140 1241 /FTID-VSP_000405.
SO	SEQUENCE 1241 AA; 137919 MW; 568544103CDF5494 CRC64;

```

Query Match      80.0%; Score 36; DB 1; Length 1241;
Best Local Similarity 77.8%; Pred. NO. 16;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Y      1 TSLVNRPP 9
b      896 TSLLKRRP 904
      ||||: |||
      ||||: |||

RESULT 9
ALBU_RABIT STANDARD; PRT; 608 AA.
P49065;
01-FEB-1996 (Rel. 33, Created)
01-FEB-1996 (Rel. 33, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Serum albumin precursor.
AUB.
Oryctolagus cuniculus (Rabbit).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
NCBI_TaxID=9986;
[1]
SEQUENCE FROM N.A.
SPRAIN-New Zealand white; TISSUE=Liver;
Sheffield W.P., Syed S., Schuyler P.D.;
Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
-1- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
-1- SUBCELLULAR LOCATION: Secreted.
-1- TISSUE SPECIFICITY: Plasma.
-1- SIMILARITY: Belongs to the ALB/AFP/VDB family.
-1- SIMILARITY: Contains 3 albumin domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch)

EMBL; U18344; AAB58347.1; -.
HSSP; P02768; 1E7B.
InterPro; IPR000264; Serum_albumin.
Pfam; PF00273; transport_prot; 3.
PRINTS; PR00802; SERUMALBUMIN.
ProDom; PD02486; Serum_albumin; 1.
SMART; SM00103; ALBUMIN; 3.
PROSITE; PS00212; ALBUMIN; 3.
Metal-binding, lipid-binding; Repeat; Signal; Copper.
SIGNAL 1 18
FT PROPER 19 24 BY SIMILARITY.
FT CHAIN 25 608 SERUM ALBUMIN.
FT DOMAIN 25 205 ALBUMIN 1.
FT DOMAIN 212 397 ALBUMIN 2.
FT DOMAIN 404 595 ALBUMIN 3.
FT METAL 27 27 COPPER.
FT METAL 27 27 BY SIMILARITY.
FT DISULFID 77 86 BY SIMILARITY.
FT DISULFID 99 115 BY SIMILARITY.
FT DISULFID 114 125 BY SIMILARITY.
FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 192 201 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 340 385 BY SIMILARITY.
FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.

```

FT DISULFID 485 501 BY SIMILARITY.
 FT DISULFID 500 511 BY SIMILARITY.
 FT DISULFID 538 583 BY SIMILARITY.
 FT DISULFID 582 591 BY SIMILARITY.
 SQ SEQUENCE 608 AA; 68914 MW; CFSE92647AAFE9A2 CRC64;
 Query Match 77.8%; Score 35; DB 1; Length 608;
 Best Local Similarity 77.8%; Pred. No. 12;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 TESLVNRRP 9
 Db 502 SESLSNRRP 510
 RESULT 10
 ATB2 OREMO STANDARD; PRT; 1112 AA.
 AC P38165;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Plasma membrane calcium-transporting ATPase 2 (EC 3.6.3.8) (PMCA2)
 DE Plasma membrane calcium pump isoform 2 (Plasma membrane calcium
 ATPase isoform 2) (Fragment).
 DE ATP2B2 OR PMCA.
 GN Oreochromis mossambicus (Mozambique tilapia) (Tilapia
 mossambica).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Labroidae;
 OC Cichlidae; Oreochromis.
 OX NCBI_TaxID=8127;
 RN SEQUENCE FROM N.A.
 RA Yang C.-H., Liu J.-H., Chou C.-M., Hwang S.-P.L., Huang C.-J.,
 RA Hwang P.-P.;
 RT Partial cDNA sequence of Mozambique tilapia (Oreochromis mossambicus)
 RL Plasma membrane calcium ATPase (PMCA).
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: This magnesium-dependent enzyme catalyzes the hydrolysis
 of ATP coupled with the transport of calcium out of the cell.
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + Ca(2+) (Cis) = ADP + phosphate +
 Ca(2+); (Trans).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to the cation transport ATPases family (P-type
 ATPases). Subfamily IIB.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF236669; AAK15034.1; -
 CC HSSP; P04191; 1EUL.
 CC InterPro; IPR006408; ATPase-IIB_Ca.
 CC InterPro; IPR001757; ATPase_EI-E2.
 CC InterPro; IPR006069; Cation_ATPase.
 CC InterPro; IPR006068; Cation_ATPase.
 CC InterPro; IPR004014; Cation_ATPase_N.
 CC InterPro; IPR008250; EI-E2_ATPase_Reg.
 CC InterPro; IPR005834; Hydrolase.
 CC Pfam; PF00689; Cation_ATPase_C; 1.
 CC Pfam; PF00690; Cation_ATPase_N; 1.
 CC Pfam; PF00122; EI-E2_ATPase; 1.
 CC Pfam; PF00702; Hydrolase; 1.
 CC PRINTS; PR00119; CATATPASE.
 CC PRINTS; PR00121; NAKATPASE.
 CC TIGRFAMs; TIGR01517; ATPase-IIB_Ca; 1.
 CC TIGRFAMs; TIGR01494; ATPase_P-type; 7.

DR PROSITE; PS00154; ATPASE_EI_E2; 1.
 KW Hydrolase; Calcium transport; Transmembrane; Phosphorylation;
 KW Magnesium; Metal-binding; ATP-binding; Calmodulin-binding.
 PT DOMAIN 1 94
 FT TRANSMEM 95 115
 FT DOMAIN 116 152
 FT TRANSMEM 153 173
 FT DOMAIN 174 373
 FT TRANSMEM 374 393
 FT DOMAIN 394 426
 FT TRANSMEM 427 444
 FT DOMAIN 445 858
 FT TRANSMEM 859 878
 FT DOMAIN 879 888
 FT TRANSMEM 889 909
 FT DOMAIN 910 929
 FT TRANSMEM 930 952
 FT DOMAIN 953 970
 FT TRANSMEM 971 992
 FT DOMAIN 993 1011
 FT TRANSMEM 1012 1033
 FT DOMAIN 1034 1043
 FT TRANSMEM 1044 1065
 FT DOMAIN 1066 >1112
 FT TRANSMEM 1106 >1112
 FT MOD_RES 482 482
 FT METAL 803 803
 FT METAL 807 807
 FT NON_TER 1112 1112
 SQ SEQUENCE 1112 AA; 122486 MW; 8BECB082E58C861 CRC64;
 Query Match 73.3%; Score 33; DB 1; Length 1112;
 Best Local Similarity 66.7%; Pred. No. 61;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 TESLVNRRP 9
 Db 914 TESLVNRRP 922
 RESULT 11
 ATB4 RAT STANDARD; PRT; 1203 AA.
 AC Q64542; Q63127; Q63445; Q64543; Q64544; Q64545;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Plasma membrane calcium-transporting ATPase 4 (EC 3.6.3.8) (PMCA4)
 DE Plasma membrane calcium pump isoform 4 (Plasma membrane calcium
 ATPase isoform 4).
 GN ATP2B4.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN SEQUENCE FROM N.A. (ISOFORMS XB; ZB; XA AND ZA).
 RC STRAIN=CD Charles River; TISSUE=Testis;
 RX MEDLINE=95217154; PubMed=7702574;
 RA Keeton T.P., Shull G.E.;
 RT "Primary structure of rat plasma membrane Ca(2+)-ATPase isoform 4 and
 RT analysis of alternative splicing patterns at splice site A.";
 RL Biochem. J. 306:779-785(1995).
 RN [2]
 RC SEQUENCE OF 266-455 FROM N.A. (ISOFORM X).
 RC TISSUE=Lung;
 RX MEDLINE=95031972; PubMed=7945253;
 RA Howard A., Barley N.F., Legon S., Walters J.R.F.;
 RT "Plasma-membrane calcium-pump isoforms in human and rat liver.";
 RL Biochem. J. 303:275-279(1994).
 RN [3]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORMS A AND B).

DOMAIN	1	92	CVTOPLASMIC (POTENTIAL).
FT	93	113	POTENTIAL.
DOMAIN	114	150	EXTRACELLULAR (POTENTIAL).
FT	151	171	POTENTIAL.
DOMAIN	172	356	CYTOPLASMIC (POTENTIAL).
FT	357	376	POTENTIAL.
DOMAIN	377	409	EXTRACELLULAR (POTENTIAL).
FT	410	427	POTENTIAL.
DOMAIN	428	840	CYTOPLASMIC (POTENTIAL).
FT	841	860	POTENTIAL.
DOMAIN	861	870	EXTRACELLULAR (POTENTIAL).
FT	871	891	POTENTIAL.
DOMAIN	892	911	CYTOPLASMIC (POTENTIAL).
FT	912	934	POTENTIAL.
DOMAIN	935	952	EXTRACELLULAR (POTENTIAL).
FT	953	974	POTENTIAL.
DOMAIN	975	993	CYTOPLASMIC (POTENTIAL).
FT	994	1015	POTENTIAL.
DOMAIN	1016	1025	EXTRACELLULAR (POTENTIAL).
FT	1026	1047	POTENTIAL.
DOMAIN	1048	1203	CYTOPLASMIC (POTENTIAL).
FT	297	303	POLY-LYS.
DOMAIN	297	303	CALMODULIN-BINDING SUBDOMAIN A
FT	1086	1103	(BY SIMILARITY).
DOMAIN	1104	1113	CALMODULIN-BINDING SUBDOMAIN B
FT	465	465	(BY SIMILARITY).
MOD RES	465	465	PHOSPHORYLATION (BY SIMILARITY).
FT	1102	1102	PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
MOD RES	785	785	MAGNESIUM (BY SIMILARITY).
FT	789	789	MAGNESIUM (BY SIMILARITY).
MOD RES	789	789	MAGNESIUM (BY SIMILARITY).
FT	301	312	Missing (in isoform ZA and isoform ZB).
MOD RES	301	312	Missing (in isoform ZA and isoform ZB).
FT	1105	1203	FTId=VSP 000406.
MOD RES	1105	1203	RVYVHFHFRDVIHAKSKNQVSIHFMPTQPEYAADMSQSF
FT	1105	1203	LNQESPSPSLAKSRITKRLSDAETVSQNNNTNNNAVDCHQVQ
MOD RES	1105	1203	INVAEFPNSPLQSQETPV -> EVINKFTQGFASFGVLRRQN
FT	1105	1203	LSQQLDVKLVPSSYGEAVASVRTSPSTSSAVTPPPVGNQSG
MOD RES	1105	1203	OSIS (in isoform XA and isoform ZA).
FT	431	431	FTId=VSP 000407.
MOD RES	431	431	I -> V (IN REF. 2).
FT	1203	1203	Score 33; DB 1; Length 1203;
MOD RES	1203	1203	Pred. No. 67;
FT	1203	1203	Mismatches 1; Indels 0; Gaps 0;
MOD RES	1203	1203	Matches 6; Conservative 2;
FT	431	431	Score 33; DB 1; Length 1203;
MOD RES	431	431	Pred. No. 67;
FT	1203	1203	Mismatches 1; Indels 0; Gaps 0;
MOD RES	1203	1203	Matches 6; Conservative 2;
FT	431	431	Score 33; DB 1; Length 1203;
MOD RES	431	431	Pred. No. 67;
FT	1203	1203	Mismatches 1; Indels 0; Gaps 0;
MOD RES	1203	1203	Matches 6; Conservative 2;
FT	431	431	Score 33; DB 1; Length 1203;
MOD RES	431	431	Pred. No. 67;
FT	1203	1203	Mismatches 1; Indels 0; Gaps 0;
MOD RES	1203	1203	Matches 6; Conservative 2;
FT	431	431	Score 33; DB 1; Length 1203;
MOD RES	431	431	Pred. No. 67;
FT	1203	1203	Mismatches 1; Indels 0; Gaps 0;
MOD RES	1203	1203	Matches 6; Conservative 2;
FT	431	431	Score 33; DB 1; Length 1203;
MOD RES	431	431	Pred. No. 67;
FT	1203	1203	Mismatches 1; Indels 0; Gaps 0;
MOD RES	1203	1203	Matches 6; Conservative 2;
FT	431	431	Score 33; DB 1; Length 1203;
MOD RES	431	431	Pred. No. 67;
FT	1203	1203	Mismatches 1; Indels 0; Gaps 0;
MOD RES	1203	1203	Matches 6; Conservative 2;
FT	431	431	Score 33; DB 1; Length 1203;
MOD RES	431	431	Pred. No. 67;
FT	1203	1203	Mismatches 1; Indels 0; Gaps 0;
MOD RES	1203	1203	Matches 6; Conservative 2;
FT	431	431	Score 33; DB 1; Length 1203;
MOD RES	431	431	Pred. No. 67;
FT	1203	1203	Mismatches 1; Indels 0; Gaps 0;
MOD RES	1203	1203	Matches 6; Conservative 2;
FT	431	431	Score 33; DB 1; Length 1203;
MOD RES	431	431	Pred. No. 67;
FT	1203	1203	Mismatches 1; Indels 0; Gaps 0;
MOD RES	1203	1203	Matches 6; Conservative 2;
FT	431	431	Score 33; DB 1; Length 1203;
MOD RES	431	431	Pred. No. 67;
FT	1203	1203	Mismatches 1; Indels 0; Gaps 0;
MOD RES	1203	1203	Matches 6; Conservative 2;
FT	431	431	Score 33; DB 1; Length 1203;
MOD RES	431	431	Pred. No. 67;
FT	1203	1203	Mismatches 1; Indels 0; Gaps 0;
MOD RES	1203	1203	Matches 6; Conservative 2;
FT	431	431	Score 33; DB 1; Length 1203;
MOD RES	431	431	Pred. No. 67;
FT	1203	1203	Mismatches 1; Indels 0; Gaps 0;
MOD RES	1203	1203	Matches 6; Conservative 2;
FT	431	431	Score 33; DB 1; Length 1203;
MOD RES	431	431	Pred. No. 67;
FT	1203	1203	Mismatches 1; Indels 0; Gaps 0;
MOD RES	1203	1203	Matches 6;

RX MEDLINE=2055576; PubMed=11105990;
RA Wang H., Ye X.Y., Ng T.B.;
RT "First demonstration of an inhibitory activity of milk proteins
RT against human immunodeficiency virus-1 reverse transcriptase and the
RT effect of succinylation."
RL Life Sci. 67:2745-2752 (2000).
CC -!- FUNCTION: Secretory RNase specific towards pyrimidine bases, with
CC higher activity towards poly C than poly U. Inhibits cell-free
CC translation. Inhibits HIV-1 reverse transcriptase.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Milk.
CC -!- SIMILARITY: Belongs to the pancreatic ribonuclease family.
DR InterPro: IPR001427; RNaseA.
DR ProDom: PD000535; RNaseA; 1.
DR PROSITE: PS00127; RNASE PANCREATIC; PARTIAL.
KW Hydrolase; Nuclease; Endonuclease; Antiviral;
KW Pyroligone carboxylic acid. PYROLIDONE CARBOXYLIC ACID.
FT MOD_RES 1 1
FT UNSURE 36 36 OR R.
FT UNSURE 37 37 OR E.
FT UNSURE 38 38 OR Q.
FT UNSURE 39 39 OR C.
FT UNSURE 40 40 OR R.
FT UNSURE 41 41 OR N.
FT UNSURE 42 42 OR C.
FT UNSURE 43 43 OR N.
FT NON_CONS 57 58
FT NON_TER 73 73
SQ SEQUENCE 73 AA; 8669 MW; 28DA862F43A8710F CRC64;

Query Match 71.1%; Score 32; DB 1; Length 73;
Best Local Similarity 75.0%; Pred. No. 5.5;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ESLVNRPP 9
DB 49 EDLTNRPP 56

RESULT 13
FLIM_BUCAI
ID FLIM_BUCAI STANDARD; PRT; 315 AA.
AC P571B2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Flagellar motor switch protein flim.
GN FLIM OR BU080.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. APS."
RL Nature 407:81-86 (2000).
CC -!- FUNCTION: FLIM IS ONE OF THREE PROTEINS (FLIG, FLIN, FLIM) THAT
CC FORM A SWITCH COMPLEX THAT IS PROPOSED TO BE LOCATED AT THE BASE
CC OF THE BASAL BODY. THIS COMPLEX INTERACTS WITH THE CHEY AND CHEZ
CC CHEMOTAXIS PROTEINS. IN ADDITION TO CONTACTING COMPONENTS OF THE
CC MOTOR THAT DETERMINE THE DIRECTION OF FLAGELLAR ROTATION (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC -!- SIMILARITY: Belongs to the flim family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP001118; BAB12800.1; -;
KW Chemotaxis; Flagellum; Membrane;
KW Complete proteome.
SQ SEQUENCE 315 AA; 36939 MW; FE5CC11D9573F198 CRC64;

Query Match 71.1%; Score 32; DB 1; Length 315;
Best Local Similarity 75.0%; Pred. No. 26;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRR 8
DB 125 TESLVNKK 132

RESULT 14
MRWV_STRMU
ID MRWV_STRMU STANDARD; PRT; 316 AA.
AC Q8DVW7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE S-adenosyl-methyltransferase mrwv (EC 2.1.1.-).
GN MRWV OR SMU.453.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar P., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
RT pathogen."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).
CC -!- FUNCTION: Exhibits a S-adenosyl-dependent methyltransferase
CC activity (By similarity).
CC -!- SIMILARITY: Belongs to the mrwv family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE014891; AAN58202.1; -;
DR HAMAP; MF 01007; -; 1.
DR InterPro: IPR002903; Bac Metrnfrse.
DR Pfam; PF01795; Methyltransf 5; 1.
DR ProDom: PD004685; Bac Metrnfrse; 1.
DR TIGRfams; TIGR00006; TIGR00006; 1.
KW Transferase; Methyltransferase; Complete proteome.
SQ SEQUENCE 316 AA; 35981 MW; 89C9D19FE5094378 CRC64;

Query Match 71.1%; Score 32; DB 1; Length 316;
Best Local Similarity 85.7%; Pred. No. 26;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 SLVNRPP 9
DB 283 SLVNRKP 289

RESULT 15

ALBU HORSE STANDARD; PRT; 607 AA.

AC P35747; 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Serum albumin precursor (Allergen Equ c 3).

GN ALE.

OS Equus caballus (Horse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

OX NCBI_TaxID=9796;

RN [1]

RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).

RC TISSUE=Liver;

RC MEDLINE=93345495; PubMed=8344282;

RA Ho J.X., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;

RT "X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm resolution."

RL Eur. J. Biochem. 215:205-212(1993).

CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Plasma.

CC -!- ALLERGEN: Causes an allergic reaction in human. Binds Ige.

CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.

CC -!- SIMILARITY: Contains 3 albumin domains.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X74045; CAA52194.1; -.

DR PIR; S24053; ABHOS.

DR HSSP; P02768; 1B7B.

DR InterPro; IPR000264; Serum_albumin.

DR Pfam; PF00273; transport_prot; 3.

DR PRINTS; PR00802; SERUMALBUMIN.

DR ProDom; PDO02486; Serum_albumin; 1.

DR SMART; SMC0103; ALBUMIN; 3.

DR PROSITE; PS00212; ALBUMIN; 3.

KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.

FT SIGNAL 1 18

FT PROPEP 19 24

FT CHAIN 25 607

FT DOMAIN 25 204

FT DOMAIN 211 396

FT DOMAIN 403 594

FT METAL 27 27

FT METAL 77 86

FT DISULFID 77 86

FT DISULFID 99 115

FT DISULFID 114 125

FT DISULFID 147 192

FT DISULFID 191 200

FT DISULFID 223 269

FT DISULFID 268 276

FT DISULFID 288 302

FT DISULFID 301 312

FT DISULFID 339 384

FT DISULFID 383 392

FT DISULFID 415 461

FT DISULFID 460 471

FT DISULFID 484 500

FT DISULFID 499 510

FT DISULFID 537 582

FT DISULFID 581 590

FT SEQUENCE 607 AA; 68598 MW; 256F6E830A1B90C5 CRC64;

Query Match 71.1%; Score 32; DB 1; Length 607;

Best Local Similarity 66.7%; Pred. No. 53;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9

DB 501 TDSLAEERP 509

Search completed: April 19, 2004, 11:52:57

Job time : 1.65651 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 3.47368 seconds
(without alignments)
817.479 Million cell updates/sec

Title: US-09-832-929-18_COPY_478_486
Perfect score: 45
Sequence: 1 TESLVNRRP 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: SPTEMBL.25.*
 - 2: sp_archaea.*
 - 3: sp_bacteria.*
 - 4: sp_fungi.*
 - 5: sp_human.*
 - 6: sp_invertebrate.*
 - 7: sp_mammal.*
 - 8: sp_mnc.*
 - 9: sp_organelle.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	395	4 Q81UK7	Q81UK7 homo sapien
2	45	100.0	417	4 Q86YGO	Q86YGO homo sapien
3	45	100.0	584	6 Q7YSG3	Q7YSG3 felis silve
4	36	80.0	1205	4 Q7Z3S1	Q7Z3S1 homo sapien
5	35	77.8	804	16 Q8EW78	Q8EW78 mycoplasma
6	35	77.8	1158	5 Q76833	Q76833 caenorhabdi
7	35	77.8	1158	5 Q9TVP9	Q9TVP9 caenorhabdi
8	35	77.8	1261	16 Q54043	Q54043 pseudomonas
9	34	75.6	283	12 Q8QTK9	Q8QTK9 banana bunc
10	34	75.6	518	5 Q8MXX1	Q8MXX1 theileria o
11	34	75.6	532	4 Q9HAQ4	Q9HAQ4 homo sapien
12	34	75.6	532	4 Q96G73	Q96G73 homo sapien
13	34	75.6	532	4 Q9GZX5	Q9GZX5 homo sapien
14	34	75.6	1281	16 Q91211	Q91211 pseudomonas
15	34	75.6	1502	5 Q81S10	Q81S10 dictyosteli
16	33	73.3	97	8 Q02658	Q02658 podospora a

17	33	73.3	323	16 Q83EJ3	Q83EJ3 coxiella bu
18	33	73.3	473	6 Q28164	Q28164 bos taurus
19	33	73.3	605	16 Q8NQY9	Q8NQY9 corynebacte
20	33	73.3	642	12 Q9DHT6	Q9DHT6 yaba-like d
21	33	73.3	849	13 Q712E6	Q712E6 brachydanio
22	33	73.3	917	10 Q9LKU3	Q9LKU3 arabidopsis
23	33	73.3	108	10 Q9ARP5	Q9ARP5 cryza sativ
24	32	71.1	146	6 Q8MJ96	Q8MJ96 equus cabal
25	32	71.1	233	4 Q9TUE2	Q9TUE2 homo sapien
26	32	71.1	289	17 Q8TRF5	Q8TRF5 methanosarc
27	32	71.1	304	16 Q81MD1	Q81MD1 bacillus an
28	32	71.1	332	16 Q8EKC2	Q8EKC2 shewanella
29	32	71.1	468	2 Q8VQS3	Q8VQS3 klebsiella
30	32	71.1	525	4 Q9H9Z3	Q9H9Z3 homo sapien
31	32	71.1	525	4 Q96176	Q96176 homo sapien
32	32	71.1	528	4 Q8NDJ2	Q8NDJ2 homo sapien
33	32	71.1	1405	5 Q9N694	Q9N694 toxoplasma
34	32	71.1	6298	11 Q8VHN7	Q8VHN7 mus musculu
35	31	68.9	94	16 Q8Y208	Q8Y208 anabaena sp
36	31	68.9	95	10 Q40248	Q40248 lactuca sat
37	31	68.9	131	16 Q82P24	Q82P24 streptomyce
38	31	68.9	154	12 Q8JKN8	Q8JKN8 heliothis z
39	31	68.9	167	16 Q8VK95	Q8VK95 mycobacteri
40	31	68.9	210	10 Q7XBQ1	Q7XBQ1 acacia hind
41	31	68.9	219	16 Q98C11	Q98C11 rhizobium l
42	31	68.9	230	16 Q92PS6	Q92PS6 rhizobium m
43	31	68.9	236	5 Q85S88	Q85S88 drosophila
44	31	68.9	237	17 Q97B18	Q97B18 thermoplasma
45	31	68.9	242	2 Q48636	Q48636 lactococcus

ALIGNMENTS

RESULT 1

Q81UK7 ID Q81UK7 PRELIMINARY; PRT; 396 AA.
AC Q81UK7
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to serum albumin precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Liver;
RC Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC035969; AAH35969.1;
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR00264; Serum albumin.
DR Pfam; PF00273; transport_prot; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;

Query Match 100.0%; Score 45; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 0.27; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

Qy 1 TESLVNRRP 9

Db 289 TESLVNRRP 297

RESULT 2

Q86YGO

```
ID Q85YGO PRELIMINARY; PRT; 417 AA.
AC Q85YGO;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to alpha-fetoprotein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041789; AA441789.1; -
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; F:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport Pfam; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR PRODOM; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 417 AA; 47360 MW; 16E764833EEF4E8D CRC64;

Query Match 100.0%; Score 45; DB 4; Length 417;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TESLVNRRP 9
Db 310 TESLVNRRP 318

RESULT 3
QYVSG3 PRELIMINARY; PRT; 584 AA.
AC QYVSG3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Albumin (Fragment).
CN ALB
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Reininger R.; Swoboda I.; Bohle B.; Hauswirth A.W.; Valent P.;
RA Rumpold H.; Valenta R.; Spitzauer S.;
RT "Escherichia coli expression and purification of recombinant cat
RT albumin: IGE recognition, induction of basophil activation and
RT lymphoproliferative responses in atopic patients.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ487677; CAD32275.1; -
DR NON_TER
FT SEQUENCE 584 AA; 65908 MW; B51002F12902C9CE CRC64;

Query Match 100.0%; Score 45; DB 6; Length 584;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TESLVNRRP 9
Db 478 TESLVNRRP 486

RESULT 4
QYVSG3
```

```
ID Q7Z3S1 PRELIMINARY; PRT; 1205 AA.
AC Q7Z3S1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein DKEZP686K08109.
GN DKEZP686K08109.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human fetal kidney;
RA Bahr A.; Lauber J.; Mewes H.W.; Weil B.; Amlid C.; Osanger A.; Fobo G.;
RA Han M.; Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX537444; CAD97686.1; -
DR XW Hypothetical protein.
SQ SEQUENCE 1205 AA; 133873 MW; 5485FCB14C243A4C CRC64;

Query Match 80.0%; Score 36; DB 4; Length 1205;
Best Local Similarity 77.8%; Pred. No. 84;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TESLVNRRP 9
Db 896 TESLVNRRP 904

RESULT 5
Q8EW78 PRELIMINARY; PRT; 804 AA.
AC Q8EW78;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cation-transporting p-type ATPase.
GN MYPE3250.
OS Mycoplasma penetrans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=28227;
RN [1]
RP SEQUENCE FROM N.A.
RC STEAIN-HF-2;
RX MEDLINE=22354719; PubMed=12466555;
RA Sasaki Y.; Ishikawa J.; Yamashita A.; Oshima K.; Kenri T.; Furuya K.;
RA Yoshino C.; Horino A.; Shiba T.; Sasaki T.; Hattori M.;
RT "The complete genomic sequence of Mycoplasma penetrans, an
RT intracellular bacterial pathogen in humans.";
RL Nucleic Acids Res. 30:5293-5300(2002).
DR EMBL; AP004171; BAC44118.1; -
DR GO; GO:0016200; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0015662; F:ATPase activity; coupled to transmembrane m. . .; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0006812; P:cation transport; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0015992; P:proton transport; IEA.
DR InterPro; IPR001757; ATPase_E1-E2.
DR InterPro; IPR008250; E1-E2_ATPase_reg.
DR InterPro; IPR005834; Hydrolase.
DR InterPro; IPR000695; H_ATPase.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR PRINTS; PR00119; CATATPASE.
DR PRINTS; PR00120; HATPASE.
DR PROSITE; PS00154; ATPASE_E1_E2; 1.
DR Complete proteome.
SQ SEQUENCE 804 AA; 88008 MW; 69C71AA628FEF7A3 CRC64;

Query Match 77.8%; Score 35; DB 16; Length 804;
Best Local Similarity 66.7%; Pred. No. 90;
```

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 TESLVNRRP 9
Db 603 TDDLNRKP 611

RESULT 6
O76833 PRELIMINARY; PRT; 1158 AA.

AC O76833
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Calcium ATPase.
GN MCA-2.

OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]

SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99134356; PubMed=9933625;
RA Kraev A.; Kaeve N.; Carafoli E.;
RT Identification and functional expression of the plasma membrane
RT calcium ATPase gene family from Caenorhabditis elegans.";
RL J. Biol. Chem. 274:4254-4258(1999).

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (E1-E2
CC ATPASES).

EMBL AJ010708; CRA09308.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0015662; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO: GO:0016787; F:hydrolase activity; IEA.
DR GO: GO:0006812; P:cation transport; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR006408; ATPase-IIB_Ca.
DR InterPro: IPR001757; ATPase_E1-E2.
DR InterPro: IPR006068; Cation_ATPase_C.
DR InterPro: IPR004014; Cation_ATPase_N.
DR InterPro: IPR008250; E1-E2_ATPase_Reg.
DR InterPro: IPR005834; Hydrolase.
DR Pfam: PF00689; Cation_ATPase_C; 1.
DR Pfam: PF00690; Cation_ATPase_N; 1.
DR Pfam: PF00122; E1-E2_ATPase; 1.
DR Pfam: PF00702; Hydrolase; 1.
DR PRINTS: PR00119; CATATPASE.
DR TIGRFAMs: TIGR01517; ATPase-IIB_Ca; 1.
DR TIGRFAMs: TIGR01494; ATPase_P-type; 6.
DR PROSITE: PS00154; ATPase_E1-E2; 1.
DR PROSITE: PS00154; ATPase_E1-E2; 1.
KW ATP-binding; Hydrolase; Phosphorylation; Transmembrane.
SQ SEQUENCE 1158 AA; 126919 MW; 18F1C35418573BCD CRC64;

Query Match 77.8%; Score 35; DB 5; Length 1158;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 TESLVNRRP 9
Db 869 TDDLNRKP 877

RESULT 7
O9TYP9 PRELIMINARY; PRT; 1158 AA.

AC O9TYP9
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN R05C11.3.

OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]

SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (E1-E2
CC ATPASES).

EMBL AF125446; AAD12806.2; -.
DR PIR: T33877; T33877.
DR WormPep: R05C11.3; CE28747.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0015662; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO: GO:0016787; F:hydrolase activity; IEA.
DR GO: GO:0006812; P:cation transport; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR006408; ATPase-IIB_Ca.
DR InterPro: IPR001757; ATPase_E1-E2.
DR InterPro: IPR006068; Cation_ATPase_C.
DR InterPro: IPR004014; Cation_ATPase_N.
DR InterPro: IPR008250; E1-E2_ATPase_Reg.
DR InterPro: IPR005834; Hydrolase.
DR Pfam: PF00689; Cation_ATPase_C; 1.
DR Pfam: PF00690; Cation_ATPase_N; 1.
DR Pfam: PF00122; E1-E2_ATPase; 1.
DR Pfam: PF00702; Hydrolase; 1.
DR PRINTS: PR00119; CATATPASE.
DR TIGRFAMs: TIGR01517; ATPase-IIB_Ca; 1.
DR TIGRFAMs: TIGR01494; ATPase_P-type; 6.
DR PROSITE: PS00154; ATPase_E1-E2; 1.
DR PROSITE: PS00154; ATPase_E1-E2; 1.
KW Hypothetical protein; ATP-binding; Hydrolase; Phosphorylation;
KW Transmembrane.
SQ SEQUENCE 1158 AA; 126920 MW; E73043B418C2CA40 CRC64;

Query Match 77.8%; Score 35; DB 5; Length 1158;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 TESLVNRRP 9
Db 869 TDDLNRKP 877

RESULT 8
O54043 PRELIMINARY; PRT; 1261 AA.

AC O54043
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Respiratory nitrate reductase alpha subunit.
GN NARG OR PA3875.
OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RA Rumpf A., Hyland S., Hoffmann T., Jahn D.;
RA Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino B., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; Y15252; CAA75540.1; -;
DR EMBL; AE004804; AAG07262.1; -;
DR PIR; G83462; G83162.
DR GO; GO:0009325; C:nitrate reductase complex; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008940; F:nitrate reductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0042126; P:nitrate metabolism; IEA.
DR InterPro; IPR009010; Asp decarb fold.
DR InterPro; IPR006656; Molybdopterin.
DR InterPro; IPR006657; Mol_dinuc_bind.
DR InterPro; IPR001005; Myb_DNA_binding.
DR InterPro; IPR006468; Narg.
DR InterPro; IPR006655; Prok_Mboxred.
DR Pfam; PF00384; molybdopterin; 1.
DR Pfam; PF01568; Molybdop_binding; 1.
DR TIGRFAWS; TIGR01580; narg; 1.
DR PROSITE; PS00551; MOLYBDOPTERIN_PROK_1; 1.
DR PROSITE; PS00490; MOLYBDOPTERIN_PROK_2; 1.
DR PROSITE; PS00932; MOLYBDOPTERIN_PROK_3; 1.
DR PROSITE; PS00037; MYB_1; 1.
KW Complete proteome.
SQ SEQUENCE 1261 AA; 140972 MW; 30CA8AF6827EB3C6 CRC64;

Query Match 77.8%; Score 35; DB 16; Length 1261;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 TESLVNRRP 9
DB 1077 TEKLNRKP 1085

RESULT 9
OQOTK9 PRELIMINARY; PRT; 283 AA.
AC OQOTK9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Replication initiation protein.
OS Banana bunchy top virus.
OC Viruses; ssDNA viruses; Nanovirus.
OX NCBI_TaxID=12585;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22034113; PubMed=12038681;
RA Bell K.E., Dale J.L., Ha C.V., Vu M.T., Revill P.A.;
RA "Characterization of Rep-encoding components associated with banana
RT bunchy top nanovirus in Vietnam.";
RL Arch. Virol. 147:695-707(2002).

DR EMBL; AF416471; AAL89631.1; -;
DR InterPro; IPR003365; Viral_rep.
DR Pfam; PF02407; Viral_Rep; 1.
SQ SEQUENCE 283 AA; 32642 MW; F4E4A22C16B718BA CRC64;
Query Match 75.6%; Score 34; DB 12; Length 283;
Best Local Similarity 87.5%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ESLVNRPP 9
DB 150 ESLVNSRP 157

RESULT 10
OQMXX1 PRELIMINARY; PRT; 518 AA.
AC OQMXX1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE P55 (Fragment).
GN TOORF2.
OS Theileria orientalis.
OC Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Theileridae;
OC Theileria.
OX NCBI_TaxID=68896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ikeda;
RA Sugimoto C., Kim J.;
RA "Novel antigenic ToORF2 gene of T.orientalis.";
RT Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AB091056; BAC10915.1; -;
DR EMBL; AB091056; BAC10915.1; -;
FT NON TER 518 518
SQ SEQUENCE 518 AA; 57016 MW; 25DFA41F5D00654 CRC64;

Query Match 75.6%; Score 34; DB 5; Length 518;
Best Local Similarity 87.5%; Pred. No. 92;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ESLVNRPP 9
DB 496 ESLVNSRP 503

RESULT 11
OQHAQ4 PRELIMINARY; PRT; 532 AA.
AC OQHAQ4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE KRAB zinc finger protein ZFOR.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21099599; PubMed=11161714;
RA Ran Q., Wadhwa R., Bischof O., Venable S., Smith J.R.,
RA Pereira-Smith O.M.;
RT "Characterization of a novel zinc finger gene with increased
RT expression in nondividing normal human cells.";
RL Exp. Cell Res. 263:156-162(2001).
DR EMBL; AF309561; AAG25714.1; -;
DR HSSP; P25490; 1UBD.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR007087; Znf_C2H2.

DR Pfam; PF01352; KRAB; 1.
 DR Pfam; PF00096; zf-C2H2; 8.
 DR PRODOM; PD000003; Znf_C2H2; 7.
 DR SMART; SM00349; KRAB; 1.
 DR SMART; SM00355; Znf_C2H2; 8.
 DR PROSITE; PS0805; KRAB; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 8.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_2; 8.
 DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 8.
 DR KW Metal-binding; Zinc; Zinc-finger.
 DR SQ SEQUENCE 532 AA; 60025 MW; E632292F95BAC8A CRC64;
 Query Match 75.6%; Score 34; DB 4; Length 532;
 Best Local Similarity 87.5%; Pred. No. 95;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TESLVNRR 8
 Db 93 SESLVNRR 100
 RESULT 12
 Q96G73 PRELIMINARY; PRT; 532 AA.
 ID Q96G73
 AC Q96G73
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Zinc-finger protein ZBRK1.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP TISSUE=Kidney;
 RA Strauberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC009921; AAH09921.1; -.
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001909; KRAB.
 DR InterPro; IPR007087; Znf_C2H2.
 DR Pfam; PF01352; KRAB; 1.
 DR Pfam; PF00096; zf-C2H2; 8.
 DR PRODOM; PD000003; Znf_C2H2; 7.
 DR SMART; SM00349; KRAB; 1.
 DR SMART; SM00355; Znf_C2H2; 8.
 DR PROSITE; PS0805; KRAB; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 8.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_2; 8.
 DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 8.
 DR KW Metal-binding; Zinc; Zinc-finger.
 DR SQ SEQUENCE 532 AA; 60011 MW; FEF12A92F95BAC8A CRC64;
 Query Match 75.6%; Score 34; DB 4; Length 532;
 Best Local Similarity 87.5%; Pred. No. 95;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TESLVNRR 8
 Db 93 SESLVNRR 100
 RESULT 13
 Q9GZX5 PRELIMINARY; PRT; 532 AA.
 ID Q9GZX5
 AC Q9GZX5
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein FLJ13405 (Zinc-finger protein ZBRK1).
 GN ZBRK1.
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP TISSUE=Placenta;
 RA Isoigai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nacai K., Sugano S., Takahashi-Fujii A., Hara H.,
 RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
 RA Arica M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
 RA Wakamatsu A., Nakamura Y., Negahari K., Masuho Y., Oshima A.,
 RT "NEBO human cDNA sequencing project";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21000486; PubMed=11090615;
 RA Zheng L., Pan H., Li S., Flesken-Nikitin A., Chen P.L., Boyer T.G.,
 RA Lee W.H.;
 RT "sequence-specific transcriptional corepressor function for BRCA1
 RT through a novel zinc-finger protein, ZBRK1.";
 RL Mol. Cell 6:757-768(2000).
 DR EMBL; AK023467; BAB14583.1; -.
 DR EMBL; AF295096; AAG17439.1; -.
 DR HSSP; P25490; IUED.
 DR Genew; HGNC:16656; ZNF350.
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001909; KRAB.
 DR InterPro; IPR007087; Znf_C2H2.
 DR Pfam; PF01352; KRAB; 1.
 DR Pfam; PF00096; zf-C2H2; 8.
 DR PRODOM; PD000003; Znf_C2H2; 7.
 DR SMART; SM00349; KRAB; 1.
 DR SMART; SM00355; Znf_C2H2; 8.
 DR PROSITE; PS0805; KRAB; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 8.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_2; 8.
 DR KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
 DR SQ SEQUENCE 532 AA; 59926 MW; 4191065B2FD2D008 CRC64;
 Query Match 75.6%; Score 34; DB 4; Length 532;
 Best Local Similarity 87.5%; Pred. No. 95;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TESLVNRR 8
 Db 93 SESLVNRR 100
 RESULT 14
 Q91211 PRELIMINARY; PRT; 1281 AA.
 ID Q91211
 AC Q91211
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein PA1923.
 GN PA1923.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PAO1;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
 RA Hickey M.J., Brinkman F.S.L., Rufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;

RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
 RT opportunistic pathogen."
 RL Nature 406:959-964(2000).
 DR EMBL: AF004618; RAG05311.1; --
 DR FIR; G83405; G83405
 DR GO; GO:0009058; P:biosynthesis; IEA.
 DR InterPro; IPR003672; COB/Mg_chitase.
 DR Pfam; PF02514; cobN-Mg_chel; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 1281 AA; 140261 MW; DB00EE98089C8C6 CRC64;

Query Match 75.6%; Score 34; DB 16; Length 1281;
 Best Local Similarity 75.0%; Pred. No. 2.5e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 ESLVNRPP 9
 Db 1269 QSVNRRP 1276

RESULT 15
 Q8IS10 PRELIMINARY; PRT; 1502 AA.
 AC Q8IS10;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 23, Last annotation update)
 DE Nucleotide exchange factor RasGEF P.
 GN GEFP.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RA Wilkins A.; Szafarski K.; Gloeckner G.; Harris Singh M.;
 RA Deenadayalan B.; Mueller R.; Eichinger L.; Noegel A.A.; Insall R.;
 RT "The family of rasGEF genes in Dictyostelium discoideum."
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY160105; ANA6895.1;
 DR GO; GO:0005885; P:guanyl-nucleotide exchange factor activity; IEA.
 DR GO; GO:0007284; P:small GTPase mediated signal transduction; IEA.
 DR InterPro; IPR001715; Calponin-like.
 DR InterPro; IPR000651; RasGEFN.
 DR InterPro; IPR001895; RasGRF CDC25.
 DR InterPro; IPR008937; Ras_GEF.
 DR Pfam; PF00307; CH; 1.
 DR Pfam; PF00617; RasGEF; 1.
 DR Pfam; PF00618; RasGEFN; 1.
 DR SMART; SM00033; CH; 1.
 DR SMART; SM00147; RasGEF; 1.
 DR SMART; SM00229; RasGEFN; 1.
 DR PROSITE; PS50021; CH; 1.
 DR PROSITE; PS50009; RasGEF CAT; 1.
 DR PROSITE; PS50212; RasGEF_NTER; 1.
 SQ SEQUENCE 1502 AA; 168915 MW; 1A53C4F11D6BF91C CRC64;

Query Match 75.6%; Score 34; DB 5; Length 1502;
 Best Local Similarity 66.7%; Pred. No. 2.9e+02;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRPP 9
 Db 996 TESLSKRP 1004

Search completed: April 19, 2004, 12:00:17
 Job time : 5.47368 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:24:29 ; Search time 5.29363 Seconds

(without alignments)
480.375 Million cell updates/sec

Title: US-09-832-929-18_COPY_478_486

Perfect score: 45

Sequence: 1 TESLVNRRP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001s:*

5: Geneseq2002s:*

6: Geneseq2003as:*

7: Geneseq2003bs:*

8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	20	7 ADB37653	Adb37653 Transglut
2	45	100.0	25	7 ADB01181	Ade01181 Bovine Se
3	45	100.0	51	2 AAY02601	Aay02601 Peptide d
4	45	100.0	73	4 AAO07138	Aao07138 Human pol
5	45	100.0	74	4 AAO00109	Aao00109 Human pol
6	45	100.0	86	4 AAU33080	AAU33080 Novel hum
7	45	100.0	119	4 AAO02645	AAO02645 Human pol
8	45	100.0	123	4 AAO04432	AAO04432 Human pol
9	45	100.0	126	4 AAU33086	AAU33086 Novel hum
10	45	100.0	133	4 AAO03606	AAO03606 Human pol
11	45	100.0	133	4 AAU33084	AAU33084 Novel hum
12	45	100.0	140	4 AAU27701	AAU27701 Human ful
13	45	100.0	151	4 AAU33085	AAU33085 Novel hum
14	45	100.0	228	3 AAY83949	Aay83949 Yeast cod
15	45	100.0	243	4 AAU33087	AAU33087 Novel hum
16	45	100.0	245	4 AAU33081	AAU33081 Novel hum
17	45	100.0	386	4 AAU33074	AAU33074 Novel hum
18	45	100.0	401	4 AAU29876	AAU29876 Novel hum
19	45	100.0	463	2 AAU14179	AAU14179 Human ser
20	45	100.0	507	4 AAU33286	AAU33286 Novel hum
21	45	100.0	507	4 AAU32994	AAU32994 Novel hum
22	45	100.0	550	4 AAU29877	AAU29877 Novel hum
23	45	100.0	582	2 AAR61713	Aar61713 Bovine se
24	45	100.0	583	2 AAW70959	AAW70959 BSA with
25	45	100.0	583	2 AAW70960	AAW70960 BSA with

26	45	100.0	583	2 AAW70961	AAW70961 BSA with
27	45	100.0	584	6 ABG72381	ABG72381 Mature hu
28	45	100.0	585	1 AAP93344	AAp93344 Sequence
29	45	100.0	585	1 AAP90388	AAp90388 Mature hu
30	45	100.0	585	1 AAP91422	AAp91422 Human nor
31	45	100.0	585	2 AAR05318	AAr05318 Human ser
32	45	100.0	585	2 AAR08457	AAr08457 Human ser
33	45	100.0	585	2 AAR26207	AAr26207 Human ser
34	45	100.0	585	2 AAR26362	AAr26362 Synthetic
35	45	100.0	585	2 AAR20029	AAr20029 Human ser
36	45	100.0	585	2 AAR80301	AAr80301 Human ser
37	45	100.0	585	2 AAO20111	AAO20111 HSA prote
38	45	100.0	585	2 AAW59841	AAW59841 Mature pr
39	45	100.0	585	3 AAY84873	AAy84873 Amino aci
40	45	100.0	585	3 AAY83946	AAy83946 Yeast cod
41	45	100.0	585	4 AAM52567	AAm52567 Mature hu
42	45	100.0	585	4 AAE12417	AAe12417 Human alb
43	45	100.0	585	4 AAE12403	AAe12403 Human alb
44	45	100.0	585	4 AAE13129	AAe13129 Human alb
45	45	100.0	585	4 AAE13135	AAe13135 Human alb

ALIGNMENTS

RESULT 1

ADB37653
ID ADB37653 standard; peptide; 20 AA.

AC ADB37653;

DT 04-DEC-2003 (first entry)

DE Transglutaminase crosslinking related peptide BSA5 SEQ ID NO:16.

KW polyvalent antigen; pAg; antigen; Ag; crosslinking; biological agent;
KW virucide; antibacterial; fungicide; antiparasitic; cytostatic; nootropic;
KW neuroprotective; vaccine; immunogenic; antigenic; medicine; infection;
KW cancer; Alzheimer's disease; immune-related disease; autoimmune disease.

OS Synthetic.

PN WO2003074004-A2.

PD 12-SEP-2003.

PF 03-MAR-2003; 2003WO-US006661.

PR 01-MAR-2002; 2002US-0361166P.

PR 08-MAR-2002; 2002US-0363445P.

PR 28-AUG-2002; 2002US-00231063.

PR 28-AUG-2002; 2002US-00231114.

PR 28-AUG-2002; 2002US-00231213.

PR 28-AUG-2002; 2002US-00231298.

PR 28-AUG-2002; 2002US-00231470.

PA (CHOU/) CHOU S.

XX Chou S;

XX WPI; 2003-756754/71.

PT Preparation of polyvalent antigen, useful in vaccines, comprises
PT crosslinking antigen in presence of biological agent, especially
PT transglutaminase, and derived antibodies.

PS Claim 140; Page 73; 130pp; English.

CC The present invention describes a method for producing a polyvalent
CC antigen (pAg). The method comprises: (i) treating an antigen (Ag) in a
CC crosslinking solution with a biological agent (I); and (ii) incubating so
CC that Ag is converted to crosslinked products. The pAg has virucide,
CC antibacterial, fungicide, antiparasitic, cytostatic, nootropic and

CC neuroprotective activities, and can be used in vaccines. The method is
CC useful in preparing immunogenic compositions, using disease-specific
CC compounds optionally modified to include a (I)-reactive amino acid that
CC are combined in crosslinking solution then treated with (I), that can be
CC used as vaccines. pAg, and other related antigenic compositions, are
CC useful in human and veterinary medicine, particularly as vaccines, for
CC treatment and prevention of infections (viral, bacterial, fungal or
CC parasitic), cancers and Alzheimer's disease, also of immune-related or
CC autoimmune diseases. Antibodies against pAg are useful as diagnostic
CC reagents and crosslinked proteins can also be used industrially, e.g. in
CC food or leather processing, in cosmetics and as enzyme carriers. The
CC present sequence represents a recombinant transglutaminase crosslinking
CC activity realted peptide, which is used in an example from the present
CC invention.

XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 45; DB 7; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
|||
Db 6 TESLVNRRP 14

RESULT 2

ADBE01181
ID ADE01181 standard; peptide; 25 AA.

XX
AC ADE01181;

XX
DT 29-JAN-2004 (first entry)

XX
DE Bovine Serum Albumin indexed peptide database peptide #11.

XX
KW mass spectrometry; peptide index; protein identification;
KW protein quantitation; protease; high-resolution mass spectrometry;
KW proteomics; Genomics; Bioinformatics; Bovine Serum Albumin.

XX
OS Bos sp.

XX
PN WO2003054549-A2.

XX
PD 03-JUL-2003.

XX
PF 09-DEC-2002; 2002WO-GB005571.

XX
PR 08-DEC-2001; 2001US-0340460P.

XX
PR 14-MAR-2002; 2002US-0364847P.

XX
PA (MICR-) MICROMASS LTD.

XX
PI Geromanos S, Dongre A, Opitack G, Silva J;

XX
DR WPI; 2003-569290/53.

XX
PT A method of mass spectrometry, useful in protein identification and
PT quantitation, by mass analyzing the first molecules in the first mixture
PT and accurately determining the mass to charge ratio of the first
PT molecules in the first mixture.

XX
PS Disclosure; Fig 7A; 134pp; English.

XX
CC The invention relates to a novel method of mass spectrometry. The method
CC comprises mass analyzing the first molecules in a first mixture and
CC accurately determining the mass to charge ratio of the first molecules in
CC the first mixture. The invention further relates to: generating an index
CC for use in identifying molecules of biological origin by mass
CC spectrometry by accurately determining the masses or mass to charge
CC ratios of molecules comprising peptides resulting from the digestion or
CC fragmentation of a polypeptide or protein; determining a first physico-
CC chemical property other than mass or mass to charge ratio of the

CC molecules comprising peptides; and optionally determining a second,
CC third, fourth and/or fifth physico-chemical property of the molecules
CC comprising peptides; and a mass spectrometer comprising a mass analyser
CC for accurately determining the mass to charge ratio of the first
CC molecules, and means for identifying the first molecules of the basis of
CC at least the first physico-chemical property and the accurately
CC determined mass to charge ratio of the first molecules and optionally on
CC the basis of the second, third, fourth and/or fifth physico-chemical
CC property. The method and spectrometer are useful in protein
CC identification, protein quantitation, proteases, high-resolution mass
CC spectrometry, proteomics, genomics and bioinformatics. This sequence
CC represents a peptide from an indexed peptide database created by the
CC novel mass spectrometry method of the invention.

XX
SQ Sequence 25 AA;

Query Match 100.0%; Score 45; DB 7; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
|||
Db 3 TESLVNRRP 11

RESULT 3

AAV02601
ID AAV02601 standard; peptide; 51 AA.

XX
AC AAV02601;

XX
DT 21-JUL-1999 (first entry)

XX
DE Peptide derived from human serum albumin.

XX
KW Compound; affinity; complementarity; human serum albumin; HSA;
KW targeted delivery.

XX
OS Homo sapiens.

XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "carboxymethyl cysteine"
FT Modified-site 2 /note= "carboxymethyl cysteine"
FT Modified-site 11 /note= "carboxymethyl cysteine"
FT Modified-site 25 /note= "carboxymethyl cysteine"
FT Modified-site /note= "LC-biotin attached"
FT Modified-site 44 /note= "carboxymethyl cysteine"

XX
PN WO9924075-A2.

XX
PD 20-MAY-1999.

XX
PF 06-NOV-1998; 98WO-US023705.

XX
PR 07-NOV-1997; 97US-0064705P.

XX
PR 13-MAR-1998; 98US-0077927P.

XX
PA (CONJ-) CONJUCHEM INC.

XX
PI Krantz A, Huang W, Hanel AM, Holmes DL, Bridon DP;

XX
DR WPI; 1999-327214/27.

XX
PT Compounds with specific affinity for human serum albumin.

XX
PS Example 12; Page 51; 78pp; English.

XX
CC The specification describes methods and compositions for identifying
CC compounds that have affinity or complementarity to a target molecule, e.g

CC with human serum albumin (HSA). These compounds are of the formula E-Ca-R
CC -Cb-A, where E is a therapeutic or diagnostic agent, R is a reactive
CC group, Ca and Cb are connector groups, and A is a pentapeptide group with
CC specific affinity for HSA. The compounds are used for targeted delivery,
CC in vivo or in vitro, or diagnostic and therapeutic agents, particularly
CC as alternatives to antibodies. The present sequence represents a peptide
CC used in the course of the invention
XX
SQ Sequence 51 AA;

Query Match 100.0%; Score 45; DB 2; Length 51;
Best Local Similarity 100.0%; Pred. No. 0.051; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

QY 1 TESLVNRRP 9
|||
Db 3 TESLVNRRP 11

RESULT 4
AAO07138
ID AAO07138 standard; protein; 73 AA.

XX AAO07138;

AC AAO07138;

DT 06-NOV-2001 (first entry)

XX Human polypeptide SEQ ID NO 21030.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

OS Homo sapiens.

XX WO200164835-A2.

PN WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004927.

XX 28-FEB-2000; 2000US-00515126.

PR 18-MAY-2000; 2000US-00577409.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Drmanac RT;

PI WPI; 2001-514838/56.

XX N-PSDB; AAI87069.

XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing

XX and treating e.g. leukemia, inflammation and immune disorders.

XX Claim 20; SEQ ID NO 21030; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation. Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 73 AA;

RESULT 6

Query Match 100.0%; Score 45; DB 4; Length 73;
Best Local Similarity 100.0%; Pred. No. 0.076; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

QY 1 TESLVNRRP 9
|||
Db 14 TESLVNRRP 22

RESULT 5
AAO00109
ID AAO00109 standard; protein; 74 AA.

XX AAO00109;

AC AAO00109;

DT 06-NOV-2001 (first entry)

XX Human polypeptide SEQ ID NO 14001.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

OS Homo sapiens.

XX WO200164835-A2.

PN WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004927.

XX 28-FEB-2000; 2000US-00515126.

PR 18-MAY-2000; 2000US-00577409.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Drmanac RT;

PI WPI; 2001-514838/56.

XX N-PSDB; AAI80040.

XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing

XX and treating e.g. leukemia, inflammation and immune disorders.

XX Claim 20; SEQ ID NO 14001; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation. Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 74 AA;

Query Match 100.0%; Score 45; DB 4; Length 74;
Best Local Similarity 100.0%; Pred. No. 0.077; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

QY 1 TESLVNRRP 9
|||
Db 40 TESLVNRRP 48

AAU33080
ID AAU33080 standard; protein; 86 AA.
AC AAU33080;
XX
DT 18-DEC-2001 (first entry)
DE
DE Novel human secreted protein #3571.
XX
XX Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 15-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-00552929.
XX
PR 26-JAN-2001; 2001US-00770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-6:11725/70.
XX
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
XX
XX Claim 20; Page 705; 765pp; English.
XX
XX The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU3304 represent the amino acid sequences of novel human
XX secreted proteins of the invention
XX
SQ Sequence 86 AA;
Query Match 100.0%; Score 45; DB 4; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.091;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TESLVNRRP 9
DB 33 TESLVNRRP 41
RESULT 7
AAO02645
ID AAO02645 standard; protein; 119 AA.
XX
AC AAO02645;
XX
XX Homo sapiens.
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 16537.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
XX
PR 18-MAY-2000; 2000US-00577409.
XX
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-514838/56.
XX
XX N-PSDB; AAI82576.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
XX Claim 20; SEQ ID NO 16537; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 119 AA;
Query Match 100.0%; Score 45; DB 4; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.13; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TESLVNRRP 9
DB 85 TESLVNRRP 93
RESULT 8
AAO04432
ID AAO04432 standard; protein; 123 AA.
XX
AC AAO04432;
XX
XX 06-NOV-2001 (first entry)
DT
XX
DE Human polypeptide SEQ ID NO 18324.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
XX WO200164835-A2.
XX
XX

PD 07-SEP-2001.
 XX
 XX 26-FEB-2001; 2001WO-US004927.
 XX
 XX 28-FEB-2000; 2000US-00515126.
 PR 18-MAY-2000; 2000US-00577409.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Drmanac RT;
 XX
 XX WPI; 2001-514838/56.
 DR N-PSDB; AA194363.
 DR
 XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
 PT and treating e.g. leukemia, inflammation and immune disorders.
 PT
 XX Claim 20; SEQ ID NO 18324; 1399pp + Sequence Listing; English.
 PS
 XX The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AA000010-AA013910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 123 AA;
 SQ

Query Match 100.0%; Score 45; DB 4; Length 123;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
 |||||
 DB 77 TESLVNRRP 85

RESULT 9
 AAU33086
 ID AAU33086 standard; protein; 126 AA.
 AC AAU33086;
 XX 18-DEC-2001 (first entry)
 DT
 XX Novel human secreted protein #3577.
 DE
 XX Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX
 XX Homo sapiens.
 OS
 XX WO200179449-A2.
 FN
 XX 25-OCT-2001.
 PD
 XX 16-APR-2001; 2001WO-US008656.
 PF
 XX 18-APR-2000; 2000US-00552929.
 PR 26-JAN-2001; 2001US-00770160.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Drmanac RT;
 XX

DR WPI; 2001-611725/70.
 XX
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy.
 PT
 XX Claim 20; Page 705-706; 765pp; English.
 PS
 XX The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used
 CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU3304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention
 XX
 XX Sequence 126 AA;
 SQ

Query Match 100.0%; Score 45; DB 4; Length 126;
 Best Local Similarity 100.0%; Pred. No. 0.14;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
 |||||
 DB 38 TESLVNRRP 46

RESULT 10
 AAU03606
 ID AAU03606 standard; protein; 133 AA.
 XX AAU03606;
 AC
 XX 06-NOV-2001 (first entry)
 DT
 XX Human polypeptide SEQ ID NO 17498.
 DE
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX
 XX Homo sapiens.
 OS
 XX WO200164835-A2.
 FN
 XX 07-SEP-2001.
 PD
 XX 26-FEB-2001; 2001WO-US004927.
 PF
 XX 28-FEB-2000; 2000US-00515126.
 PR 18-MAY-2000; 2000US-00577409.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Drmanac RT;
 PI
 XX WPI; 2001-514838/56.
 DR N-PSDB; AA183537.
 DR
 XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
 PT and treating e.g. leukemia, inflammation and immune disorders.
 PT
 XX Claim 20; SEQ ID NO 17498; 1399pp + Sequence Listing; English.
 PS

XX The invention relates to human polynucleotides (AA179941-AA193941) and
 CC the encoded proteins (AA00010-AA013910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 133 AA;

Query Match 100.0%; Score 45; DB 4; Length 133;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
 |||||
 Db 117 TESLVNRRP 125

RESULT 11
 AAU33084
 ID AAU33084 standard; protein; 133 AA.

AC AAU33084;

DT 18-DEC-2001 (first entry)

DE Novel human secreted protein #3575.

XX Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

OS Homo sapiens.

PN WO200179449-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001WO-US008656.

PR 18-APR-2000; 2000US-00552929.

PR 26-JAN-2001; 2001US-00770160.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-611725/70.

PT Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy.

PS Claim 20; Page 705; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used

CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU3304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention
 XX
 XX Sequence 133 AA;

Query Match 100.0%; Score 45; DB 4; Length 133;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
 |||||
 Db 33 TESLVNRRP 41

RESULT 12
 AAU27701
 ID AAU27701 standard; protein; 140 AA.

AC AAU27701;

DT 18-DEC-2001 (first entry)

DE Human full-length polypeptide sequence #26.

XX Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
 KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
 KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
 KW nervous system disorder; inflammatory disorder; cell differentiation;
 KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
 KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
 KW cytostatic; antirheumatic; antiarthritic; vulnery; antiparkinsonian;
 KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
 KW neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
 KW immunostimulant; analgesic; gene therapy.

OS Homo sapiens.

PN WO200164834-A2.

PD 07-SEP-2001.

PF 26-FEB-2001; 2001WO-US004926.

PR 28-FEB-2000; 2000US-00515126.

PR 18-MAY-2000; 2000US-00577403.

PR 17-JUN-2000; 2000US-00597707.

PR 14-JUL-2000; 2000US-00616807.

PR 19-SEP-2000; 2000US-00664641.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
 PI Drmanac R;

DR WPI; 2001-589862/66.

DR N-PSDB; AAS44601.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis, treatment of cancer,
 PT neurological, inflammatory disorders and for use in arrays for detection.

PS Claim 10; SEQ ID NO 198; 153pp; English.

XX Sequences AAU27676-AAU28019 represent full-length polypeptides and contig
 CC polypeptides of the invention. The proteins and their associated DNA
 CC sequences are useful for the treatment, diagnosis and prevention of
 CC various types of disorder in a mammalian subject such as a human, dog,
 CC monkey, mouse, hamster or rat. The disorders include cancers such as

CC Leukaemia, lymphoma and neuroblastoma, autoimmune disorders such as
CC multiple sclerosis, connective tissue disease, rheumatoid arthritis,
CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's
CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
CC bowel disease. The sequences exhibit activity relating to angiogenesis,
CC cell proliferation, cell differentiation, stem cell growth factor,
CC activin or inhibin. Therefore, they can be used to manipulate stem cells
CC in culture to give rise to neuroepithelial cells that can be used to
CC augment or replace cells damaged by illness, accidental damage or genetic
CC disorders. The sequences may also be used for regeneration of bone,
CC cartilage, tendons and ligaments and in tissue repair and burn healing.
CC Note: Some sequences for this patent did not form part of the printed
CC specification, but were obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 140 AA;

Query Match 100.0%; Score 45; DB 4; Length 140;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
Db 33 TESLVNRRP 41
|||||

RESULT 13
AAU33085
ID AAU33085 standard; protein; 151 AA.
XX
AC AAU33085;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #3576.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-00552929.
XX
PR 26-JAN-2001; 2001US-00770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
XX
PS Claim 20; Page 705; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the

CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU3304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 151 AA;

Query Match 100.0%; Score 45; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
Db 51 TESLVNRRP 59
|||||

RESULT 14
AAU83949
ID AAU83949 standard; protein; 228 AA.
XX
AC AAU83949;
XX
DT 28-JUL-2000 (first entry)
XX
DE Yeast codon-biased recombinant HSA protein fragment HSA-III.
XX
KW Recombinant; human serum albumin; HSA; yeast codon bias; host cell;
KW overlapping oligonucleotide; expression vector.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN CN1239103-A.
XX
PD 22-DEC-1999.
XX
PF 17-JUN-1998; 98CN-00102506.
XX
PR 17-JUN-1998; 98CN-00102506.
XX
PA (HAIJ-) HAIJI BIOENGINEERING CO LTD.
XX
PI Li S, Lu D;
XX
WPI; 2000-351138/31.
XX
DR N-PSDB; AAA10094.
XX
PT Process for preparing recombinant human serum albumin comprising yeast
PT biased sex codons - uses a recombinant DNA technique.
XX
PS Example 1; Fig 7; 44pp; Chinese.
XX
CC The method relates to a method of recombinantly producing human serum
CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise
CC a yeast codon bias. The complete HSA gene (AAA10091) was generated as
CC three synthetic fragments (AAA10092-AA10094) joined by recombinant DNA
CC technology. Each HSA fragment was synthesised from overlapping
CC oligonucleotide fragments that were extended. This sequence represents
CC the sequence of the HSA fragment HSA-III encoded by the human gene with a
CC yeast codon bias. The invention also covers a recombinant expression
CC vector, yeast host cells carrying the recombinant expression vector and
CC the process for producing human serum albumin in the yeast host cell,
CC especially in secretory mode
XX
SQ Sequence 228 AA;

Query Match 100.0%; Score 45; DB 3; Length 228;
Best Local Similarity 100.0%; Pred. No. 0.27;

Job time : 6.29363 secs

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TESLVNRRP 9
Db 121 TESLVNRRP 129

RESULT 15
AAU33087
ID AAU33087 standard; protein; 243 AA.

AC AAU33087;
XX
XX 18-DEC-2001 (first entry)
XX
XX Novel human secreted protein #3578.
XX
XX Human; vaccination; gene therapy; nutritional supplement;
XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;
XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.

OS Homo sapiens.
XX
XX WO200179449-A2.
XX
XX 25-OCT-2001.
XX
XX 16-APR-2001; 2001WO-US008656.
XX
XX 18-APR-2000; 2000US-00552929.
XX
XX 26-JAN-2001; 2001US-00770160.
XX
XX (HYSE-) HYSEQ INC.

XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-611725/70.
XX
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy.

XX Claim 20; Page 706; 765pp; English.
XX
XX The invention relates to novel human secreted polypeptides. The
XX polypeptides and antibodies to the polypeptides are useful for
XX determining the presence of or predisposition to a disease associated
XX with altered levels of polypeptide. The polypeptides are also useful for
XX identifying agents (agonists and antagonists) that bind to them. Cells
XX expressing the proteins are useful for identifying a therapeutic agent
XX for use in treatment of a pathology related to aberrant expression or
XX physiological interactions of the polypeptide. Vectors comprising the
XX nucleic acids encoding the polypeptides and cells genetically engineered
XX to express them are also useful for producing the proteins. The proteins
XX are useful in genetic vaccination, testing and therapy, and can be used
XX as nutritional supplements. They may be used to increase stem cell
XX proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
XX and/or nerve tissue growth or regeneration; immune suppression and/or
XX stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
XX AAU29510-AAU3304 represent the amino acid sequences of novel human
XX secreted proteins of the invention

SQ Sequence 243 AA;

Query Match 100.0%; Score 45; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TESLVNRRP 9
Db 138 TESLVNRRP 146

Search completed: April 19, 2004, 11:51:22

OM protein - protein search, using sw model

Run on: April 19, 2004, 12:00:25 ; Search time 3.78947 Seconds
(without alignments)
654.724 Million cell updates/sec

Title: US-09-832-929-18_COPY_478_486

Perfect score: 45

Sequence: 1 TESLVNRRP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/FCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	20	12	US-10-231-063C-16
2	45	100.0	20	15	US-10-231-298B-16
3	45	100.0	20	15	US-10-231-470C-16
4	45	100.0	20	15	US-10-231-213D-16
5	45	100.0	20	15	US-10-231-114C-16
6	45	100.0	84	12	US-10-424-599-259972
7	45	100.0	583	15	US-10-360-101-200
8	45	100.0	585	9	US-09-929-552-2
9	45	100.0	585	10	US-09-932-613-445
10	45	100.0	585	10	US-09-984-010-26
11	45	100.0	585	10	US-09-833-041-18
12	45	100.0	585	10	US-09-833-117-18
13	45	100.0	585	10	US-09-932-323-445
14	45	100.0	585	10	US-09-832-501-18
15	45	100.0	585	11	US-09-833-118-18

16	45	100.0	585	11	US-09-833-245-18	Sequence 18, Appl
17	45	100.0	585	12	US-10-424-999-11	Sequence 11, Appl
18	45	100.0	585	12	US-10-425-000-31	Sequence 31, Appl
19	45	100.0	585	12	US-10-433-108-34	Sequence 34, Appl
20	45	100.0	585	13	US-10-153-064-5	Sequence 5, Appl
21	45	100.0	585	14	US-10-153-604A-5	Sequence 5, Appl
22	45	100.0	585	14	US-10-319-263-1	Sequence 1, Appl
23	45	100.0	585	14	US-10-319-263-2	Sequence 2, Appl
24	45	100.0	585	14	US-10-414-469-1	Sequence 1, Appl
25	45	100.0	585	14	US-10-414-469-2	Sequence 2, Appl
26	45	100.0	585	14	US-10-413-831-1	Sequence 1, Appl
27	45	100.0	585	14	US-10-413-831-2	Sequence 2, Appl
28	45	100.0	585	15	US-10-413-832-1	Sequence 1, Appl
29	45	100.0	585	15	US-10-413-832-2	Sequence 2, Appl
30	45	100.0	585	15	US-10-414-386-1	Sequence 1, Appl
31	45	100.0	585	15	US-10-414-386-2	Sequence 2, Appl
32	45	100.0	585	15	US-10-233-675A-11	Sequence 11, Appl
33	45	100.0	585	15	US-10-462-262-26	Sequence 26, Appl
34	45	100.0	604	10	US-09-984-010-7	Sequence 7, Appl
35	45	100.0	604	13	US-10-045-170A-1	Sequence 1, Appl
36	45	100.0	607	14	US-10-057-789-45	Sequence 45, Appl
37	45	100.0	607	14	US-10-212-628-45	Sequence 45, Appl
38	45	100.0	609	10	US-09-919-039-370	Sequence 370, App
39	45	100.0	609	12	US-10-609-346-12	Sequence 12, Appl
40	45	100.0	609	13	US-10-153-064-7	Sequence 7, Appl
41	45	100.0	609	14	US-10-153-604A-7	Sequence 7, Appl
42	45	100.0	609	14	US-10-365-623-23	Sequence 23, Appl
43	45	100.0	610	9	US-09-984-186-2	Sequence 2, Appl
44	45	100.0	610	14	US-10-237-667-2	Sequence 2, Appl
45	45	100.0	610	14	US-10-237-708-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-10-231-063C-16
; Sequence 16, Application US/10231063C
; Publication No. US2003022476A1
; GENERAL INFORMATION:
; APPLICANT: Chou, Szu-Yi
; TITLE OF INVENTION: Method of Producing Transglutaminase Reactive Compound
; FILE REFERENCE: SAMG/0004
; CURRENT APPLICATION NUMBER: US/10/231,063C
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 60/361,166
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/363,445
; PRIOR FILING DATE: 2002-03-08
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; TYPE: PRT
; LENGTH: 20
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic Bovine Serum Albumin peptide
US-10-231-063C-16

Query Match 100.0%; Score 45; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TESLVNRRP 9
| | | | |
Db 6 TESLVNRRP 14

RESULT 2
US-10-231-298B-16
; Sequence 16, Application US/10231298B
; Publication No. US20030219853A1
; GENERAL INFORMATION:

; APPLICANT: Chou, Szu-Yi
; TITLE OF INVENTION: Method of Cross-Linking a Compound
; FILE REFERENCE: SAMG/0006
; CURRENT APPLICATION NUMBER: US/10/231,298B
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 60/361,166
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/363,445
; PRIOR FILING DATE: 2002-03-08
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic Bovine Serum Albumin peptide
US-10-231-298B-16

Query Match 100.0%; Score 45; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
DB 6 TESLVNRRP 14

RESULT 3
US-10-231-470C-16
; Sequence 16, Application US/10231470C
; Publication No. US20030219857A1
; GENERAL INFORMATION:
; APPLICANT: Chou, Szu-Yi
; TITLE OF INVENTION: Method of Producing Transglutaminase Having Broad Substrate
; FILE REFERENCE: SAMG/0003
; CURRENT APPLICATION NUMBER: US/10/231,470C
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 60/361,166
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/363,445
; PRIOR FILING DATE: 2002-03-08
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic Bovine Serum Albumin peptide
US-10-231-470C-16

Query Match 100.0%; Score 45; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
DB 6 TESLVNRRP 14

RESULT 4
US-10-231-213D-16
; Sequence 16, Application US/10231213D
; Publication No. US20040001848A1
; GENERAL INFORMATION:
; APPLICANT: Chou, Szu-Yi
; TITLE OF INVENTION: Method of Producing Disease-Specific Antigens
; FILE REFERENCE: SAMG/0005
; CURRENT APPLICATION NUMBER: US/10/231,213D
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 60/361,166

; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/363,445
; PRIOR FILING DATE: 2002-03-08
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic Bovine Serum Albumin peptide
US-10-231-213D-16

Query Match 100.0%; Score 45; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
DB 6 TESLVNRRP 14

RESULT 5
US-10-231-114C-16
; Sequence 16, Application US/10231114C
; Publication No. US20040005654A1
; GENERAL INFORMATION:
; APPLICANT: Chou, Szu-Yi
; TITLE OF INVENTION: Method of Producing Polyvalent Antigens
; FILE REFERENCE: SAMG/0002
; CURRENT APPLICATION NUMBER: US/10/231,114C
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 60/361,166
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/363,445
; PRIOR FILING DATE: 2002-03-08
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic Bovine Serum Albumin peptide
US-10-231-114C-16

Query Match 100.0%; Score 45; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
DB 6 TESLVNRRP 14

RESULT 6
US-10-424-599-259972
; Sequence 259972, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 259972
; LENGTH: 84
; TYPE: PRT

; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(84)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_7677C.1.pcp
US-10-424-599-259972

Query Match 100.0%; Score 45; DB 12; Length 84;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
Db 22 TESLVNRRP 30

RESULT 7
US-10-360-101-200
; Sequence 200, Application US/10360101
; Publication No. US2004000950A1
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N.
; APPLICANT: Leenhouts, Cornelis J.
; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360,101
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: EP 02077060.8
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 200
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of albumin
US-10-360-101-200

Query Match 100.0%; Score 45; DB 15; Length 583;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
Db 477 TESLVNRRP 485

RESULT 8
US-09-929-552-2
; Sequence 2, Application US/09929552
; Patent No. US20020123080A1
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/929,552

; FILING DATE: 14-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/769,746
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-929-552-2

Query Match 100.0%; Score 45; DB 9; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
Db 478 TESLVNRRP 486

RESULT 9
US-09-932-613-445
; Sequence 445, Application US/09932613
; Publication No. US20030091565A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON
; FILE REFERENCE: DYX-025.1 PCT; DYX-025.1 US
; CURRENT APPLICATION NUMBER: US/09/932,613
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: HomoSapiens
US-09-932-613-445

Query Match 100.0%; Score 45; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TESLVNRRP 9
Db 478 TESLVNRRP 486

RESULT 10
US-09-984-010-26
; Sequence 26, Application US/09984010
; Publication No. US20030104578A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David James
; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
; STREET: 1300 I Street, NW